Gray-Wyner and Slepian-Wolf Guessing

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

***Abstract*—We study the guessing variants of two distributed source coding problems: the Gray-Wyner network and the Slepian-Wolf network. Building on the former, we propose a new definition of the Rényi common information as the least attainable common rate in the Gray-Wyner guessing problem under the no-excess-rate constraint. We then provide a variational characterization of this quantity. In the Slepian-Wolf setting, we follow up the work of Bracher-Lapidoth-Pfister with the case where the expected number of guesses need not converge to one but must be dominated by some given exponential.**

I. Introduction and Problem Statement

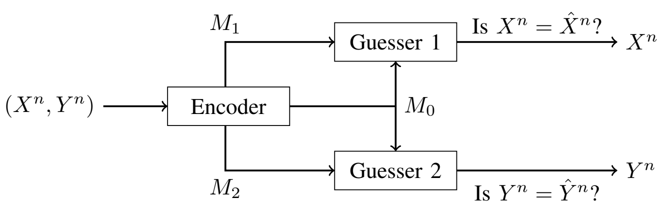

Fig. 1. Gray-Wyner Guessing Setup

Gray-Wyner guessing. A length-n sequence $(X^n, Y^n) \triangleq (X_1, Y_1), \ldots, (X_n, Y_n)$ of tuples is drawn IID according to a PMF P_{XY} on the finite set $\mathcal{X} \times \mathcal{Y}$. A rate (R_0, R_1, R_2)-encoder $\phi = (\phi_0, \phi_1, \phi_2)$

$$\begin{aligned} \phi\colon \mathcal{X}^n \times \mathcal{Y}^n &\to \{0,1\}^{nR_0} \times \{0,1\}^{nR_1} \times \{0,1\}^{nR_2} \\ (x^n, y^n) &\mapsto \big(\phi_0(x^n, y^n), \phi_1(x^n, y^n), \phi_2(x^n, y^n)\big) \end{aligned} \tag{1}$$

describes the sequence (X^n, Y^n) as $(M_0, M_1, M_2) \triangleq \phi(X^n, Y^n)$. The pair (M_0, M_1) is revealed to Guesser 1, who wishes to recover X^n, and the pair (M_0, M_2) to Guesser 2, who wishes to recover Y^n (see Fig. 1). To recover X^n, Guesser 1—after observing (M_0, M_1)—chooses a guessing order

$$\operatorname{ord}_X\colon \{1, \ldots, |\mathcal{X}|^n\} \xrightarrow{\text{bijection}} \mathcal{X}^n \tag{2}$$

on $\mathcal{X}^n$, and guesses

"Is $X^n = \operatorname{ord}_X(1)$?", "Is $X^n = \operatorname{ord}_X(2)$?", ...

until correct (and X^n hence revealed). The number of guesses taken by Guesser 1 is denoted $G_X(X^n)$, with G_X being the inverse function of ord_X, i.e., $\operatorname{ord}_X(G_X(x^n)) = x^n$ for all $x^n \in \mathcal{X}^n$. To recover Y^n, Guesser 2 proceeds analogously, with the guessing order on $\mathcal{Y}^n$ and its inverse function denoted ord_Y and G_Y. Note that, while the encoder ϕ and the guessing orders ord_X and ord_Y depend on n, we do not make this dependence explicit; n will typically be clear from the context.

Given $\rho > 0$ and a sequence of encoders and guessing orders, we define the guessing exponents

$$E_X \triangleq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}[G_X(X^n)^\rho] \tag{3a}$$

$$E_Y \triangleq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}[G_Y(Y^n)^\rho]. \tag{3b}$$

We say that a rate tuple $(R_0, R_1, R_2) \in \mathbb{R}^3_{\geq 0}$ is $(\bar{E}_X, \bar{E}_Y)$-achievable in the Gray-Wyner guessing problem if for every $\epsilon > 0$ there exists a sequence of encoders ϕ and guessing orders $\operatorname{ord}_X, \operatorname{ord}_Y$ for which

$$E_X \leq \bar{E}_X + \epsilon \text{ and } E_Y \leq \bar{E}_Y + \epsilon. \tag{4}$$

We denote the set of all $(\bar{E}_X, \bar{E}_Y)$-achievable rate tuples $\mathcal{R}^\rho_{GW}(\bar{E}_X, \bar{E}_Y)$, with the shorthand exception $\mathcal{R}^\rho_{GW} \triangleq \mathcal{R}^\rho_{GW}(0,0)$. In Sections II and III we characterize $\mathcal{R}^\rho_{GW}(\bar{E}_X, \bar{E}_Y)$ as follows:

Theorem 1. *When $(X^n, Y^n) \sim \text{IID}\, P_{XY}$, the Gray-Wyner guessing region $\mathcal{R}^\rho_{GW}(\bar{E}_X, \bar{E}_Y)$ equals*

$$\begin{aligned} \bigcap_{Q_{XY}} \Bigg(\bigcup_{Q_{U|XY}} \Big\{ &(R_0, R_1, R_2) \in \mathbb{R}^3_{\geq 0}\colon R_0 \geq \mathrm{I}_Q(U; X, Y), \\ &R_1 \geq \mathrm{H}_Q(X \mid U) - \frac{1}{\rho}(\mathrm{D}(Q_{XY} \| P_{XY}) + \bar{E}_X), \\ &R_2 \geq \mathrm{H}_Q(Y \mid U) - \frac{1}{\rho}(\mathrm{D}(Q_{XY} \| P_{XY}) + \bar{E}_Y) \Big\} \Bigg), \end{aligned} \tag{5}$$

where U can be any chance variable whose support $\mathcal{U}$ is finite; the intersection is over all PMFs Q_{XY} on $\mathcal{X} \times \mathcal{Y}$; the union is over all conditional PMFs $Q_{U|XY}$ on $\mathcal{U} \times \mathcal{X} \times \mathcal{Y}$; and H_Q and I_Q denote the entropy and mutual information computed w.r.t. $Q_{XY} Q_{U|XY}$.

Note that substituting 0 for $\bar{E}_X$ and $\bar{E}_Y$ in (5) and replacing the intersection over Q_{XY} with the substitution of P_{XY} for Q_{XY} yields the set of achievable rates in the Gray-Wyner source coding problem [1, Theorem 4].

For $\bar{E}_X = \bar{E}_Y = 0$, we define the least achievable sum-rate as

$$R^*_{\rho,\Sigma} \triangleq \inf \Big\{ R_0 + R_1 + R_2\colon (R_0, R_1, R_2) \in \mathcal{R}^\rho_{GW} \Big\}. \tag{6}$$

The following variational characterization of $R^*_{\rho,\Sigma}$ is provided without proof.

Theorem 2. *When* $(X^n, Y^n) \sim \mathrm{IID}\, P_{XY}$,

$$R^*_{\rho,\Sigma} = \sup_{Q_{XY}} \inf_{\substack{Q_{U|XY}:\ \mathrm{H}_Q(X|U) \le \mathrm{D}(Q_{XY}\|P_{XY})/\rho \\ \mathrm{H}_Q(Y|U) \le \mathrm{D}(Q_{XY}\|P_{XY})/\rho}} \mathrm{I}_Q(U;X,Y), \tag{7}$$

where U can be any chance variable whose support $\mathcal{U}$ is finite; the supremum is over all PMFs Q_{XY} on $\mathcal{X}\times\mathcal{Y}$; and the infimum is over all conditionals PMFs $Q_{U|XY}$ on $\mathcal{U}\times\mathcal{X}\times\mathcal{Y}$.

Note that when Y^n is deterministic, $R^*_{\rho,\Sigma}$ equals the order-$1/(1+\rho)$ Rényi entropy $\mathrm{H}_{1/(1+\rho)}(X)$ of X (cf. [2, Proposition 8] for the variational characterization of the Rényi entropy).

Following Wyner's arguments in [3], we propose the following operational definition of the *Rényi common information* of order $1/(1+\rho)$ between X and Y:

$$\begin{aligned}&\mathrm{C}_{1/(1+\rho)}(X;Y)\\ &\triangleq \inf\left\{R_0 \ge 0 \colon \exists (R_1,R_2) \text{ s.t. } \begin{matrix}(R_0,R_1,R_2) \in \mathcal{R}^\rho_{GW} \\ R_0+R_1+R_2 = R^*_{\rho,\Sigma}\end{matrix}\right\}.\end{aligned} \tag{8}$$

Combining Theorem 1 and 2, we obtain the following variational characterization of $\mathrm{C}_{1/(1+\rho)}(X;Y)$:

Theorem 3. *The order-$1/(1+\rho)$ Rényi common information $\mathrm{C}_{1/(1+\rho)}(X;Y)$ corresponding to the joint PMF P_{XY} is*

$$\sup_{Q_{XY}} \inf_{\substack{Q_{U|XY}:\ (\mathrm{H}_Q(X|U)-\mathrm{D}(Q_{XY}\|P_{XY})/\rho)^+ \\ +\,(\mathrm{H}_Q(Y|U)-\mathrm{D}(Q_{XY}\|P_{XY})/\rho)^+ \\ +\,\mathrm{I}_Q(U;X,Y) \le R^*_{\rho,\Sigma}}} \mathrm{I}_Q(U;X,Y), \tag{9}$$

where $(x)^+ \triangleq \max(x,0)$; where U can be any chance variable whose support $\mathcal{U}$ is finite; the supremum is over all PMFs Q_{XY} on $\mathcal{X}\times\mathcal{Y}$; and the infimum is over all conditional PMFs $Q_{U|XY}$ on $\mathcal{U}\times\mathcal{X}\times\mathcal{Y}$.

Alternative definitions of a Rényi counterpart to Wyner's common information have been proposed in [4] and [5].

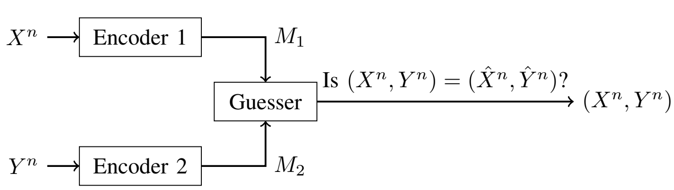

Fig. 2. Slepian-Wolf Guessing Setup

Slepian-Wolf guessing. Let $(X^n,Y^n) \sim \mathrm{IID}\, P_{XY}$. A rate-$R_1$ encoder ϕ_1 for X^n and a rate-R_2 encoder ϕ_2 for Y^n

$$\phi_1 \colon \mathcal{X}^n \to \{0,1\}^{nR_1}, \quad \phi_2 \colon \mathcal{Y}^n \to \{0,1\}^{nR_2}, \tag{10}$$

describe (X^n,Y^n) as $(M_1,M_2) \triangleq (\phi_1(X^n),\phi_2(Y^n))$. The pair (M_1,M_2) is revealed to a guesser who wishes to recover both X^n and Y^n (see Fig. 2). To do so, the guesser—after observing (M_1,M_2)—fixes a guessing order

$$\mathrm{ord}_{XY} \colon \{1,\ldots,|\mathcal{X}\times\mathcal{Y}|^n\} \xrightarrow{\text{bijection}} \mathcal{X}^n \times \mathcal{Y}^n \tag{11}$$

on $\mathcal{X}^n\times\mathcal{Y}^n$, and guesses

"Is $(X^n,Y^n) = \mathrm{ord}_{XY}(1)$?",
"Is $(X^n,Y^n) = \mathrm{ord}_{XY}(2)$?",
...

until correct. Analogously to the Gray-Wyner setting, we denote the number of guesses by $G_{XY}(X^n,Y^n)$, with G_{XY} being the inverse function of ord_{XY}.

Given $\rho > 0$ and a sequence of encoders and guessing orders, we define the guessing exponents

$$E_{XY} \triangleq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}[G_{XY}(X^n,Y^n)^\rho]. \tag{12}$$

A rate tuple $(R_1,R_2) \in \mathbb{R}^2_{\ge 0}$ is $\bar{E}_{XY}$-achievable in the Slepian-Wolf guessing problem if for every $\epsilon > 0$ there exist a sequence of encoders ϕ_1, ϕ_2 and guessing orders ord_{XY} for which $E_{XY} \le \bar{E}_{XY} + \epsilon$. The set of all achievable rate tuples is denoted $\mathcal{R}^\rho_{SW}(\bar{E}_{XY})$. In Sections IV and V we prove the following characterization of $\mathcal{R}^\rho_{SW}(\bar{E}_{XY})$:

Theorem 4. *When $(X^n,Y^n) \sim \mathrm{IID}\, P_{XY}$, the Slepian-Wolf guessing region $\mathcal{R}^\rho_{SW}(\bar{E}_{XY})$ equals*

$$\begin{aligned}\bigcap_{Q_{XY}} \Big\{&(R_1,R_2) \in \mathbb{R}^2_{\ge 0} \colon \\ &R_1 \ge \mathrm{H}_Q(X \mid Y) - \frac{1}{\rho}(\mathrm{D}(Q_{XY}\|P_{XY}) + \bar{E}_{XY}), \\ &R_2 \ge \mathrm{H}_Q(Y \mid X) - \frac{1}{\rho}(\mathrm{D}(Q_{XY}\|P_{XY}) + \bar{E}_{XY}), \\ &R_1+R_2 \ge \mathrm{H}_Q(X,Y) - \frac{1}{\rho}(\mathrm{D}(Q_{XY}\|P_{XY}) + \bar{E}_{XY})\Big\},\end{aligned} \tag{13}$$

where the intersection is over all PMFs Q_{XY} on $\mathcal{X}\times\mathcal{Y}$.

Note that substituting 0 for $\bar{E}_{XY}$ in (13) and replacing the intersection over Q_{XY} with the substitution of P_{XY} for Q_{XY} yields the set of achievable rates in the Slepian-Wolf source coding problem [6]. Further observe that, using the variational definition of the Rényi entropy, (13) can be simplified as follows:

Corollary 1. *The Slepian-Wolf guessing region $\mathcal{R}^\rho_{SW}(\bar{E}_{XY})$ equals the set of all $(R_1,R_2) \in \mathbb{R}^2_{\ge 0}$ satisfying*

$$R_1 \ge \mathrm{H}_{1/(1+\rho)}(X \mid Y) - \frac{1}{\rho}\bar{E}_{XY} \tag{14a}$$

$$R_2 \ge \mathrm{H}_{1/(1+\rho)}(Y \mid X) - \frac{1}{\rho}\bar{E}_{XY} \tag{14b}$$

$$R_1+R_2 \ge \mathrm{H}_{1/(1+\rho)}(X,Y) - \frac{1}{\rho}\bar{E}_{XY}, \tag{14c}$$

where $H_{1/(1+\rho)}(\cdot \mid \cdot)$ is the conditional Rényi entropy of order $1/(1+\rho)$.

Note that, by [7, Theorem 1], for $\bar{E}_{XY} = 0$ any tuple (R_1,R_2) satisfying (14) with strict inequalities is also achievable in the stronger sense that there exists a sequence of encoders and guessing orders for which, not only is E_{XY} zero, but $\limsup_{n\to\infty} \mathbb{E}[G_{XY}(X^n,Y^n)^\rho] = 1$.

II. Gray-Wyner Guessing, Achievability

Below we prove the direct part of Theorem 1, namely, that for fixed $\rho > 0$, $\bar{E}_X \geq 0$, $\bar{E}_Y \geq 0$, and (R_0, R_1, R_2) in (5), there exist for every $\epsilon > 0$ a sequence of rate-(R_0, R_1, R_2) encoders ϕ and guessing orders ord_X, ord_Y for which $E_X \leq \bar{E}_X + \epsilon$ and $E_Y \leq \bar{E}_Y + \epsilon$. Throughout the proof we assume $R_0 > 0$, because otherwise (5) implies $R_1 \geq \mathrm{H}_{1/(1+\rho)}(X) - \bar{E}_X/\rho$ and $R_2 \geq \mathrm{H}_{1/(1+\rho)}(Y) - \bar{E}_Y/\rho$, which is achievable by describing X^n and Y^n separately [8, Eq. (5)].

We begin by introducing some notation: for a fixed positive integer n, let $\mathcal{P}^n(\mathcal{X} \times \mathcal{Y})$ denote the set of denominator-n types on $\mathcal{X} \times \mathcal{Y}$, i.e., the set of rational PMFs on $\mathcal{X} \times \mathcal{Y}$ with denominator n. For $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$, let $Q_{x^n y^n}$ or $\hat{Q}_{XY}$ (when x^n and y^n are clear from the context) denote the empirical joint type of (x^n, y^n),

$$Q_{x^n y^n}(x, y) = \frac{1}{n} \mathrm{N}(x, y \mid x^n, y^n), \tag{15}$$

where $\mathrm{N}(x, y \mid x^n, y^n)$ is the number of occurrences of (x, y) in (x^n, y^n). And for $Q \in \mathcal{P}^n(\mathcal{X} \times \mathcal{Y})$, let $\mathcal{T}^n(Q)$ denote the type class of Q, i.e., the set of all $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ whose empirical type is Q.

To prove the direct part of Theorem 1, we proceed as follows: Given ρ, $\bar{E}_X$, $\bar{E}_Y$, (R_0, R_1, R_2) as above and an arbitrary $\epsilon > 0$, we will construct a sequence of rate-(R_0, R_1, R_2) encoders ϕ and guessing orders ord_X and ord_Y for which $E_X \leq \bar{E}_X + \epsilon$ and $E_Y \leq \bar{E}_Y + \epsilon$. Our construction will be based on the following two observations: 1) Because $|\mathcal{P}^n(\mathcal{X} \times \mathcal{Y})|$ grows only polynomially in n [9, Theorem 11.1.1], the encoder can describe the joint type $\hat{Q}_{XY}$ of (X^n, Y^n) as part of the common message M_0 without increasing its rate. 2) For every $\epsilon'' > 0$ there exists a positive integer $n_{\epsilon''}$, such that for all $n \geq n_{\epsilon''}$ the following holds: For every $Q_{XY} \in \mathcal{P}^n(\mathcal{X} \times \mathcal{Y})$ and every conditional type $Q_{U|XY}$ (for which $Q \triangleq Q_{XY} Q_{U|XY}$ is in $\mathcal{P}^n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$), there exists a codebook $\mathcal{C}(Q_{XY}, Q_{U|XY}) \subseteq \mathcal{U}^n$, later denoted $\mathcal{C}(Q_{XY})$, whose size is at most $2^{n(\mathrm{I}_Q(U;X,Y) + \epsilon'')}$ and satisfying that for every $(x^n, y^n) \in \mathcal{T}^n(Q_{XY})$ there is some codeword $u^n \in \mathcal{C}(Q_{XY}, Q_{U|XY})$ with $(x^n, y^n, u^n) \in \mathcal{T}^n(Q)$. This fact is sometimes referred to as the Type Covering Lemma [10, Lemma 2.34].

Using the above observations, we construct an encoder ϕ as follows: First, we fix ϵ' and ϵ'' sufficiently small and $n \geq n_{\epsilon''}$ sufficiently large (how to choose ϵ', ϵ'', and n will become apparent later in the proof). Every $Q_{XY} \in \mathcal{P}^n(\mathcal{X} \times \mathcal{Y})$ we map to a conditional type $Q_{U|XY}(Q_{XY})$ satisfying

$$R_0 \geq \mathrm{I}_Q(U; X, Y) + \epsilon' + \epsilon'' \tag{16a}$$

$$R_1 \geq \mathrm{H}_Q(X \mid U) - \frac{1}{\rho}(\mathrm{D}(Q_{XY} \| P_{XY}) + \bar{E}_X + \epsilon) \tag{16b}$$

$$R_2 \geq \mathrm{H}_Q(Y \mid U) - \frac{1}{\rho}(\mathrm{D}(Q_{XY} \| P_{XY}) + \bar{E}_Y + \epsilon), \tag{16c}$$

where $Q = Q_{XY} Q_{U|XY}(Q_{XY})$. Such a conditional type exists because (R_0, R_1, R_2) lies by assumption in (5) and because every PMF can be approximated arbitrary well by a type of sufficiently large denominator n. Our chosen mapping $Q_{XY} \mapsto Q_{U|XY}(Q_{XY})$ is revealed to the encoder and guessers. For every $Q_{XY} \in \mathcal{P}^n(\mathcal{X} \times \mathcal{Y})$, let $\mathcal{C}(Q_{XY})$ be the codebook whose existence is guaranteed by Observation 2. Reveal this codebook to all parties. Finally, for every $Q_{XY} \in \mathcal{P}^n(\mathcal{X} \times \mathcal{Y})$ and $u^n \in \mathcal{C}(Q_{XY})$, we partition the conditional type class $\mathcal{T}^n(Q_{X|U}|u^n)$ (i.e., the set of all $x^n \in \mathcal{T}^n(Q_X)$ for which x^n and u^n are jointly typical w.r.t Q_{XU}, where Q_X and Q_{XU} are the X-marginal and (X, U)-marginal of Q) into 2^{nR_1} equally sized bins, and $\mathcal{T}^n(Q_{Y|U}|u^n)$ into 2^{nR_2} equally sized bins. Reveal these partitions to all parties. The encoder can now be described as follows: To describe (X^n, Y^n), it uses the first $n\epsilon'$ bits of the common message M_0 to describe the empirical joint type $\hat{Q}_{XY}$ of (X^n, Y^n). It then uses the remaining $n(R_0 - \epsilon')$ bits of M_0 to describe some $U^n \in \mathcal{C}(\hat{Q}_{XY})$ that is jointly typical with (X^n, Y^n) w.r.t. $\hat{Q}_{XY} Q_{U|XY}(\hat{Q}_{XY})$. Finally, the encoder uses message M_1 to describe the bin of $\mathcal{T}^n(\hat{Q}_{X|U}|U^n)$ containing X^n, and message M_2 to describe the bin of $\mathcal{T}^n(\hat{Q}_{Y|U}|U^n)$ containing Y^n.

From the first $n\epsilon'$ bits of M_0 both guessers recover $\hat{Q}_{XY}$; from $\hat{Q}_{XY}$ they recover the conditional type $Q_{U|XY}(\hat{Q}_{XY})$ and the codebook $\mathcal{C}(\hat{Q}_{XY})$; and from $\mathcal{C}(\hat{Q}_{XY})$ and the last $n(R_0 - \epsilon')$ bits of M_0 they recover U^n. Knowing U^n and the empirical joint type $\hat{Q}_{XU}$ of (X^n, U^n), Guesser 1 picks an arbitrary guessing order on the elements of the bin indexed by M_1, namely,

$$\{x^n \in \mathcal{T}^n(\hat{Q}_{X|U}|U^n) \colon \phi_1(x^n) = M_1\} \tag{17}$$

(ignoring all x^n not belonging to (17)). Guesser 2 proceeds analogously, picking an arbitrary guessing order on the set of all $y^n \in \mathcal{T}^n(\hat{Q}_{Y|U}|U^n)$ assigned to M_2.

We next analyze the proposed guessing scheme, beginning with an upper bound on $\mathbb{E}[G_X(X^n)^\rho]$. Denoting conditional expectation given the event $\{\hat{Q}_{XY} = Q_{XY}\}$ by $\mathbb{E}_{Q_{XY}}$,

$$\mathbb{E}[G_X(X^n)^\rho] \tag{18}$$

$$\overset{(a)}{=} \sum_{Q_{XY}} \mathbb{E}_{Q_{XY}}[G_X(X^n)^\rho] \, \mathbb{P}[\hat{Q}_{XY} = Q_{XY}] \tag{19}$$

$$\overset{(b)}{\leq} \sum_{Q_{XY}} \mathbb{E}_{Q_{XY}}[G_X(X^n)^\rho] 2^{-n \mathrm{D}(Q_{XY} \| P_{XY})} \tag{20}$$

$$\overset{(c)}{\leq} \sum_{Q_{XY}} \Big(\big| \{x^n \in \mathcal{T}^n(Q_{X|U}|U^n) \colon \phi_1(x^n) = M_1\} \big|^\rho \; 2^{-n \mathrm{D}(Q_{XY} \| P_{XY})} \Big) \tag{21}$$

$$\overset{(d)}{=} \sum_{Q_{XY}} 2^{n(\bar{E}_X + \epsilon + \mathrm{D}(Q_{XY} \| P_{XY}))} 2^{-n \mathrm{D}(Q_{XY} \| P_{XY})} \tag{22}$$

$$\overset{(e)}{=} 2^{n(\bar{E}_X + \epsilon + \delta_n)}, \tag{23}$$

where (a) follows from the law of total expectation; (b) is due to Sanov's Theorem [9, Theorem 11.4.1]; (c) holds because the guesser will at most try every element from (17) (with $\hat{Q}_{XU} = Q_{XU}$); (d) holds because the conditional type class $\mathcal{T}^n(Q_{X|U}|U^n)$ contains at most $2^{n\,\mathrm{H}_Q(X|U)}$ elements, because it is evenly partitioned into 2^{nR_1} bins, and because $R_1 \geq \mathrm{H}_Q(X \mid U) - \frac{1}{\rho}(\mathrm{D}(Q_{XY}\|P_{XY}) + \bar{E}_X + \epsilon)$; and (e) follows from the fact that the number of denominator-n types on $\mathcal{X} \times \mathcal{Y}$ grows polynomially in n, and where δ_n is hence a sequence tending to zero as n tends to infinity. From (23) we see that

$$E_X = \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}[G_X(X^n)^\rho] \leq \bar{E}_X + \epsilon, \tag{24}$$

and by adapting (19)–(23) to $\mathbb{E}[G_Y(Y^n)^\rho]$,

$$E_Y = \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}[G_Y(Y^n)^\rho] \leq \bar{E}_Y + \epsilon, \tag{25}$$

which completes the proof of the direct part of Theorem 1.

III. Gray-Wyner Guessing, Converse

We now prove the converse part of Theorem 1. Fix $\rho > 0$, $\bar{E}_X \geq 0$, $\bar{E}_Y \geq 0$, and $(R_0, R_1, R_2) \in \mathbb{R}^3_{\geq 0}$. We will show that if for every $\epsilon > 0$ there exists a sequence of rate-(R_0, R_1, R_2) encoders ϕ and guessing orders ord_X, ord_Y for which $E_X \leq \bar{E}_X + \epsilon$ and $E_Y \leq \bar{E}_Y + \epsilon$, then for every PMF $\tilde{Q}_{XY}$ on $\mathcal{X} \times \mathcal{Y}$ there exists a conditional PMF $Q_{U|XY}$ such that

$$R_0 \geq \mathrm{I}_{\tilde{Q}}(U; X, Y) \tag{26a}$$

$$R_1 \geq \mathrm{H}_{\tilde{Q}}(X \mid U) - \frac{1}{\rho}(\mathrm{D}(\tilde{Q}_{XY}\|P_{XY}) + \bar{E}_X) \tag{26b}$$

$$R_2 \geq \mathrm{H}_{\tilde{Q}}(Y \mid U) - \frac{1}{\rho}(\mathrm{D}(\tilde{Q}_{XY}\|P_{XY}) + \bar{E}_Y), \tag{26c}$$

where $\tilde{Q} = \tilde{Q}_{XY} Q_{U|XY}$. To show this, fix a rate tuple (R_0, R_1, R_2) and $\epsilon > 0$, and consider a sequence of encoders ϕ (of these rates) and guessing orders ord_X, ord_Y satisfying $E_X \leq \bar{E}_X + \epsilon/2$ and $E_Y \leq \bar{E}_Y + \epsilon/2$. Starting with (19) and this time invoking the lower bound bound in Sanov's Theorem, we obtain

$$\mathbb{E}[G_X(X^n)^\rho] \geq \sum_{Q_{XY}} \mathbb{E}_{Q_{XY}}[G_X(X^n)^\rho] 2^{-n(\mathrm{D}(Q_{XY}\|P_{XY}) - \delta_n)}, \tag{27}$$

where the sum is over all $Q_{XY} \in \mathcal{P}^n(\mathcal{X} \times \mathcal{Y})$. Define

$$E_X(Q_{XY}) \triangleq \frac{1}{n} \log \mathbb{E}_{Q_{XY}}[G_X(X^n)^\rho]. \tag{28}$$

The assumption $E_X \leq \bar{E}_X + \epsilon/2$ and (27) imply that for all large enough n and all $Q_{XY} \in \mathcal{P}^n(\mathcal{X} \times \mathcal{Y})$,

$$E_X(Q_{XY}) - \mathrm{D}(Q_{XY}\|P_{XY}) \leq \bar{E}_X + \epsilon. \tag{29}$$

Analogously,

$$E_Y(Q_{XY}) - \mathrm{D}(Q_{XY}\|P_{XY}) \leq \bar{E}_Y + \epsilon, \tag{30}$$

where $E_Y(Q_{XY}) \triangleq \frac{1}{n} \log \mathbb{E}_{Q_{XY}}[G_Y(Y^n)^\rho]$.

Next we show that since (29) and (30) hold for every $\epsilon > 0$, (26) must hold. To that end, first note that by [11, Theorem 1] and the fact that the Rényi entropy $\mathrm{H}_\alpha(\cdot)$ is non-increasing in α,

$$E_{Q_{XY}}[G_X(X^n)^\rho] \geq 2^{\rho\,\mathrm{H}_Q(X^n|M_0,M_1) - n\delta_n}, \tag{31}$$

where Q is the joint law of $(X^n, Y^n, M_0, M_1, M_2)$ conditioned on the event $\{\hat{Q}_{XY} = Q_{XY}\}$,

$$\begin{aligned} &Q(x^n, y^n, m_0, m_1, m_2) \\ &= \frac{1}{|\mathcal{T}^n(Q_{XY})|} \cdot \begin{cases} 1, & \text{if } Q_{x^n y^n} = Q_{XY} \\ & \text{and } (m_0, m_1, m_2) = \phi(x^n, y^n) \\ 0, & \text{else.} \end{cases} \end{aligned} \tag{32}$$

We next lower-bound $\mathrm{H}_Q(X^n \mid M_0, M_1)$ as follows:

$$\begin{aligned} \mathrm{H}_Q(X^n \mid M_0, M_1) &\overset{(a)}{\geq} \mathrm{H}_Q(X^n \mid M_0) - nR_1 &(33)\\ &= \sum_{i=1}^n \mathrm{H}_Q(X_i \mid X^{i-1}, M_0) - nR_1 &(34)\\ &\overset{(b)}{\geq} \sum_{i=1}^n \mathrm{H}_Q(X_i \mid U_i) - nR_1 &(35)\\ &\overset{(c)}{=} n(\mathrm{H}_Q(X_T \mid U_T, T) - R_1) &(36)\\ &\overset{(d)}{=} n(\mathrm{H}_{\tilde{Q}}(X \mid U) - R_1), &(37)\end{aligned}$$

where (a) holds because M_1 can assume at most 2^{nR_1} distinct values; in (b) we have conditioned on Y^{i-1} (in addition to (X^{i-1}, M_0)) and defined the chance variable $U_i \triangleq (X^{i-1}, Y^{i-1}, M_0)$ taking values in $\mathcal{U}_i \triangleq \mathcal{X}^{i-1} \times \mathcal{Y}^{i-1} \times \{0,1\}^{nR_0}$; in (c) we have introduced the chance variable T that is uniform over $\{1, \ldots, n\}$ and independent of $(X^n, Y^n, M_0, M_1, M_2)$ (and implicitly extended the domain of Q to include T); and in (d) we have defined $U \triangleq (U_T, T)$ and the PMF $\tilde{Q}$ on $\mathcal{X} \times \mathcal{Y} \times \left(\cup_{i=1}^n \mathcal{U}_i\right) \times \{1, \ldots, n\}$:

$$\tilde{Q}(x, y, u, t) = \frac{1}{n} \mathbb{P}[(X_t, Y_t, U_t) = (x, y, u)], \tag{38}$$

where the probability on the RHS is computed w.r.t. Q. Recall that under Q, (X^n, Y^n) is uniform over $\mathcal{T}^n(Q_{XY})$, and thus the (X,Y)-marginal $\tilde{Q}_{XY}$ of $\tilde{Q}$ equals Q_{XY},

$$\tilde{Q}_{XY}(x, y) = Q_{XY}(x, y), \quad \forall (x^n, y^n) \in \mathcal{T}^n(Q_{XY}). \tag{39}$$

Combining (39), (37), (31), (29), and (28), we obtain a lower bound on R_1:

$$R_1 \geq \mathrm{H}_{\tilde{Q}}(X \mid U) - \frac{1}{\rho}(\mathrm{D}(\tilde{Q}_{XY}\|P_{XY}) + \bar{E}_X) - \frac{\epsilon}{\rho} - \delta_n. \tag{40}$$

Analogously,

$$R_2 \geq \mathrm{H}_{\tilde{Q}}(Y \mid U) - \frac{1}{\rho}(\mathrm{D}(\tilde{Q}_{XY}\|P_{XY}) + \bar{E}_Y) - \frac{\epsilon}{\rho} - \delta_n. \tag{41}$$

Having established that (29) and (30) imply (40) and (41), whose right-hand sides (RHSs) approach those of (26b) and (26c), we next consider R_0.

$$\begin{aligned} nR_0 &\geq \mathrm{H}_Q(M_0) &(42)\\ &\geq \mathrm{I}_Q(X^n, Y^n; M_0) &(43)\\ &= \mathrm{H}_Q(X^n, Y^n) - \mathrm{H}_Q(X^n, Y^n \mid M_0) &(44)\end{aligned}$$

$$\stackrel{(a)}{\geq} n(\mathrm{H}_{\tilde{Q}}(X,Y)-\delta_n) - \mathrm{H}_Q(X^n,Y^n \mid M_0) \tag{45}$$

$$= n(\mathrm{H}_{\tilde{Q}}(X,Y)-\delta_n) - \sum_{i=1}^{n} \mathrm{H}_Q(X_i,Y_i \mid U_i) \tag{46}$$

$$= n(\mathrm{I}_{\tilde{Q}}(U;X,Y)-\delta_n), \tag{47}$$

where (a) holds because under Q, (X^n,Y^n) is uniform over $\mathcal{T}^n(Q_{XY})$ and because $|\mathcal{T}^n(Q_{XY})| \geq 2^{n(\mathrm{H}_{\tilde{Q}}(X,Y)-\delta_n)}$.

We now observe that the RHSs of (40), (41), and (47) approach those of (26) as we let ϵ tend to zero and n to infinity. Note that while $\tilde{Q}_{XY}$ in (40), (41), and (47) is a type, (26) nevertheless holds for arbitrary PMFs because the RHS of (26) is continuous in $\tilde{Q}_{XY}$, and because any PMF can be approximated arbitrarily well by an appropriate type of sufficiently large denominator. This concludes the proof of the converse part of Theorem 1.

IV. Slepian-Wolf Guessing, Achievability

We now prove the direct part of Theorem 4. Fix $\rho > 0$ and $\bar{E}_{XY} \geq 0$. We show that for every $\epsilon > 0$ and (R_1,R_2) in (13), there exists a sequence of rate-R_1 encoders ϕ_1, rate-R_2 encoders ϕ_2 and guessing orders ord_{XY} for which $E_{XY} \leq \bar{E}_{XY}+\epsilon$. We prove the existence of those using a random binning argument: For every $Q_X \in \mathcal{P}^n(\mathcal{X})$, we assign every $x^n \in \mathcal{T}^n(Q_X)$ a random index $M^1_{Q_X}(x^n) \in \{1,\ldots,2^{nR_1}\}$ (chosen independently and uniformly), and likewise for every $Q_Y \in \mathcal{P}^n(\mathcal{Y})$, we assign every $y^n \in \mathcal{T}^n(Q_Y)$ a random index $M^2_{Q_Y}(y^n) \in \{1,\ldots,2^{nR_2}\}$. The assignments $x^n \mapsto M^1_{Q_{x^n}}(x^n)$ and $y^n \mapsto M^1_{Q_{y^n}}(y^n)$ are revealed to the encoders and guesser. The message M_1 produced by Encoder 1 is $M^1_{\hat{Q}_X}(X^n)$, and the message M_2 produced by Encoder 2 is $M^2_{\hat{Q}_Y}(Y^n)$. To construct the guessing order ord_{XY}, we use the Interlaced Guessing Lemma [2, Proposition 6] which, for the purpose of this proof, asserts that we may assume that the guesser is cognizant of the empirical joint type $\hat{Q}_{XY}$ of (X^n,Y^n). Under this assumption, the guesser chooses an arbitrary guessing order on the set

$$\mathcal{G}(X^n,Y^n) \triangleq \Big\{(\xi^n,\eta^n) \in \mathcal{T}^n(\hat{Q}_{XY})\colon M^1_{\hat{Q}_X}(\xi^n)=M_1, \\ M^2_{\hat{Q}_Y}(\eta^n)=M_2\Big\} \tag{48}$$

(all (ξ^n,η^n) not belonging to (48) are ignored). We next examine the proposed guessing scheme. In the following all probabilities and expectations are computed over X^n, Y^n, and the binning. For lack of space, some of the arguments are abbreviated.

Because the number of guesses is at most the number of elements in (48), our goal is to upper-bound $\mathbb{E}[|\mathcal{G}(X^n,Y^n)|^\rho]$. By the law of total expectation and Sanov's Theorem,

$$\mathbb{E}[|\mathcal{G}(X^n,Y^n)|^\rho] \leq \sum_{Q_{XY}} \Big(\mathbb{E}_{Q_{XY}}[|\mathcal{G}(X^n,Y^n)|^\rho] \\ 2^{-n\,\mathrm{D}(Q_{XY}\|P_{XY})}\Big), \tag{49}$$

where $\mathbb{E}_{Q_{XY}}$ denotes expectation (over X^n, Y^n, and the binning) conditioned on the event $\{\hat{Q}_{XY}=Q_{XY}\}$. For every $Q_{XY} \in \mathcal{P}^n(\mathcal{X}\times\mathcal{Y})$ and $(\xi^n,\eta^n) \in \mathcal{T}^n(Q_{XY})$, let $Z_{Q_{XY}}(\xi^n,\eta^n)$ be one if $(\xi^n,\eta^n) \in \mathcal{G}(X^n,Y^n)$, and zero otherwise (whether it is one or zero hence depends on X^n, Y^n, and the random mapping to the bins). Observe that

$$\mathbb{E}_{Q_{XY}}[|\mathcal{G}(X^n,Y^n)|^\rho] = \mathbb{E}_{Q_{XY}}\left[\left(\sum_{(\xi^n,\eta^n)} Z_{Q_{XY}}(\xi^n,\eta^n)\right)^\rho\right], \tag{50}$$

where the sum is over all $(\xi^n,\eta^n) \in \mathcal{T}^n(Q_{XY})$. Denoting the conditional probability measure given the event $\{\hat{Q}_{XY}=Q_{XY}\}$ by $\mathbb{P}_{Q_{XY}}$, one can show that for every $(\xi^n,\eta^n) \in \mathcal{T}^n(Q_{XY})$,

$$\mathbb{P}_{Q_{XY}}[Z_{Q_{XY}}(\xi^n,\eta^n)=1] \leq 2^{-n(\mathrm{H}_Q(X,Y)-\delta_n)} + 2^{-n(R_1+R_2)} \\ + 2^{-nR_1}2^{-n(\mathrm{H}_Q(Y)-\delta_n)} + 2^{-nR_2}2^{-n(\mathrm{H}_Q(X)-\delta_n)}. \tag{51}$$

After using this to upper-bound the RHS of (50), one can further show that

$$\mathbb{E}_{Q_{XY}}[|\mathcal{G}(X^n,Y^n)|^\rho] \leq \Big(1+2^{n\rho(H_Q(X|Y)-R_1)} \\ + 2^{n\rho(H_Q(Y|X)-R_2)} + 2^{n\rho(\mathrm{H}_Q(X,Y)-R_1-R_2)}\Big)2^{n\delta_n}. \tag{52}$$

From (52) and (49),

$$\mathbb{E}_{Q_{XY}}[|\mathcal{G}(X^n,Y^n)|^\rho] \leq 2^{n(\bar{E}_{XY}+\delta'_n)}, \tag{53}$$

where the sequence $\{\delta'_n(\delta_n)\}_{n\in\mathbb{N}}$ tends to zero as n tends to infinity. Taking the logarithm and letting $n\to\infty$ on both sides in (53) concludes the proof of the direct part of Theorem 4.

V. Slepian-Wolf Guessing, Converse

Omitted.

References

[1] R. Gray and A. Wyner, "Source coding for a simple network," *Bell Syst. Tech. J.*, vol. 53, pp. 1681 – 1721, Nov. 1974.

[2] A. Lapidoth, "Lecture notes on guessing and Rényi entropy," 2019.

[3] A. Wyner, "The common information of two dependent random variables," *IEEE Trans. Inf. Theory*, vol. 21, pp. 163 – 179, Mar. 1975.

[4] L. Yu and V. Tan, "Wyner's common information under Rényi divergence measures," *IEEE Trans. Inf. Theory*, vol. 64, pp. 3616 – 3632, Feb. 2018.

[5] ——, "On exact and ∞-Rényi common informations," in *Int. Symp. Inf. Theory*, Jul. 2019, pp. 2229 – 2233.

[6] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, pp. 471 – 480, Jul. 1973.

[7] A. Bracher, A. Lapidoth, and C. Pfister, "Guessing with distributed encoders," *Entropy*, vol. 21, no. 3, 2019.

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

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

[10] S. Moser, *Advanced topics in information theory (lecture notes)*, 4th ed., 2019. [Online]. Available: http://moser-isi.ethz.ch/scripts.html

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