Source Coding for Synthesizing Correlated Randomness

Touheed Anwar Atif
University of Michigan, USA
Email: touheed@umich.edu

Arun Padakandla
University of Tennessee, USA
Email: arunpr@utk.edu

S. Sandeep Pradhan
University of Michigan, USA
Email: pradhanv@umich.edu

***Abstract*—We consider a scenario wherein two parties Alice and Bob are provided X_1^n and X_2^n – samples that are IID from a PMF $p_{X_1X_2}$. Alice and Bob can communicate to Charles over (noiseless) communication links of rate R_1 and R_2 respectively. Their goal is to enable Charles generate samples Y^n such that the triple (X_1^n, X_2^n, Y^n) has a PMF that is close, in total variation, to $\prod p_{X_1X_2Y}$. In addition, the three parties may posses shared common randomness at rate C. We address the problem of characterizing the set of rate triples (R_1, R_2, C) for which the above goal can be accomplished. We provide a set of sufficient conditions, i.e., an achievable rate region for this three party setup. Our work also provides a complete characterization of a point-to-point setup wherein Bob is absent and Charles is provided with side-information.**

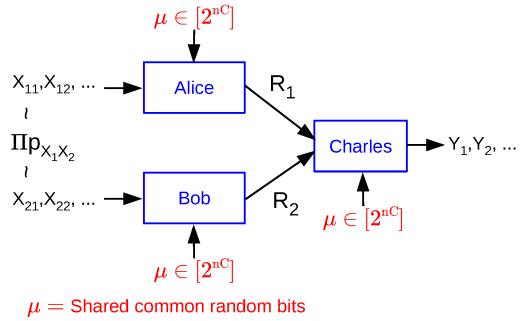

Fig. 1. Source Coding for Synthesizing Correlated Randomness

I. Introduction

The task of generating correlated randomness at different terminals in a network has applications in several communication and computing scenarios. This task also serves as a primitive in several cryptographic protocols. In this article, we study the problem of characterizing fundamental information-theoretic limits of generating such correlated randomness in network scenarios.

We consider the scenario depicted in Fig 1. Three distributed parties - Alice, Bob and Charles - have to generate samples that are independent and identically distributed (IID) with a target probability mass function (PMF) $p_{X_1X_2Y}$. Alice and Bob are provided with samples that are IID $p_{X_1X_2}$ - the corresponding marginal of the target PMF $p_{X_1X_2Y}$. They have access to unlimited private randomness and share noiseless communication links of rates R_1, R_2 with Charles. In addition, the three parties share common randomness at rate C. For what rate triples (R_1, R_2, C) can Alice and Bob enable Charles to generate the required samples? In this article, we provide a set of sufficient conditions, i.e., an achievable rate region. In the process, we provide an alternate solution for the two-terminal version wherein Charles is provided with side-information and Bob is absent. This problem also stems out as a special case from the problem setup described in [1], however, the technique used here facilitates in characterizing the rate-region for the scenario in Fig 1.

The problem of characterizing communication rates required to generate correlated randomness at distributed terminals can be traced back to the work of Wyner [2]. Wyner considered the scenario of distributed parties generating IID samples distributed with PMF p_{XY}, when fed with a common information stream. In characterizing the minimum rate of this common information stream, Wyner discovered a fundamental tool - the technique of *soft covering*. Soft covering has found applications in diverse areas including computer science, classical and quantum information theory. As we illustrate in the sequel, this work adds another dimension to our current understanding of soft covering.

A renewed interest in soft covering led Cuff [3] to consider a point-to-point (PTP) version of the scenario depicted in Fig. 1, wherein Bob (or X_2) is absent. Leveraging [2], [4], Cuff [3] provided a characterization of the minimum rate R_1 for all values of the common randomness rate C. Cuff's work shares an interesting connection with an analogous problem in quantum information. Prior to [3], Winter [5] considered the problem of simulating quantum measurements with limited common randomness. This work was generalized in [6] where the authors characterized a complete trade-off between communication and common randomness rates. Building on this, [7] studied a distributed scenario consisting of three distributed parties and derived inner and outer bounds.

Cuff's [3] findings rely on the use of a likelihood encoder that maps the observed sequence and common random bits into a codebook of sufficient rate. Essentially, the encoder performs a MAP decoding of the observed sequence into the chosen codebook. While this choice greatly simplifies the analysis, it permits little room for generalization. Our experience in network information theory suggests that encoding and decoding via joint-typicality can be naturally generalized to diverse multi-terminal scenarios. Motivated by this, we propose joint-typicality based encoding and decoding to perform soft covering. In view of the general applicability of typicality-based coding schemes, we regard the typicality-based soft covering we propose as an important step.

II. Preliminaries and Problem Statement

We supplement standard information theory notation with the following. For a PMF p_X, we let $p_X^n = \prod_{i=1}^n p_X$. For an integer $n \geq 1$, $[n] := \{1, \cdots, n\}$. The total variation between

This work was supported by NSF grant CCF 1717299.

PMFs p_X and q_X defined over $\mathcal{X}$ is denoted $\|p_X - q_X\|_1 = \frac{1}{2}\sum_{x\in\mathcal{X}} |p_X(x) - q_X(x)|$.

Definition 1. *Given a PMF p_{XYZ}, a rate pair (R, C) is achievable, if $\forall \epsilon > 0$ and sufficiently large n, there exists a collection of 2^{nC} randomized encoders $E^{(\mu)} : \mathcal{X}^n \to [\Theta]$ for $\mu \in [2^{nC}]$ and a corresponding collection of 2^{nC} randomized decoders $D^{(\mu)} : \mathcal{Z}^n \times [\Theta] \to \mathcal{Y}^n$ for $\mu \in [2^{nC}]$ such that $|p^n_{XYZ} - p_{X^nY^nZ^n}|_1 \le \epsilon$, $\frac{1}{n}\log_2 \Theta \le R + \epsilon$, where*

$$p_{X^nY^nZ^n}(x^n, y^n, z^n) = \sum_{\mu\in[2^{nC}]} 2^{-nC} \sum_{m\in[\Theta]} p^n_{XZ}(x^n, z^n) \, p^{(\mu)}_{M|X^n}(m|x^n) p^{(\mu)}_{Y^n|Z^n,M}(y^n|z^n, m),$$

$p^{(\mu)}_{M|X^n}, p^{(\mu)}_{Y^n|Z^nM}$ are the PMFs induced by encoder and decoder respectively, corresponding to shared random message μ. We let $\mathcal{R}_s(p_{XYZ})$ denote the set of achievable rate pairs.

Cuff [3, Thm. II.1] provides a characterization for $\mathcal{R}_s(p_{XY})$ when $\mathcal{Z} = \phi$ is empty. Our first main result (Thm. 1) is a characterization of $\mathcal{R}_s(p_{XYZ})$. Building on this, we address the network scenario (Fig. 1) for which we state the problem below. In the following, we let $\underline{X} = (X_1, X_2), \underline{x}^n = (x_1^n, x_2^n)$.

Definition 2. *Given a PMF $p_{X_1X_2Y}$, a rate triple (R_1, R_2, C) is achievable, if $\forall \epsilon > 0$ and sufficiently large n, there exists 2^{nC} randomized encoder pairs $E^{(\mu)}_j : \mathcal{X}^n_j \to [\Theta_j] : j \in [2], \mu \in [2^{nC}]$, and a corresponding collection of 2^{nC} randomized decoders $D^{(\mu)} : [\Theta_1] \times [\Theta_2] \to \mathcal{Y}^n$ for $\mu \in [2^{nC}]$ such that $\left|p^n_{\underline{X}Y} - p_{\underline{X}^nY^n}\right|_1 \le \epsilon$, $\frac{1}{n}\log_2 \Theta_j \le R_j + \epsilon : j \in [2]$, where*

$$p_{\underline{X}^nY^n}(\underline{x}^n, y^n) = \sum_{\mu\in[2^{nC}]} 2^{-nC} \sum_{\substack{(m_1,m_2)\in\\ [\Theta_1]\times[\Theta_2]}} p^n_{\underline{X}Y}(\underline{x}^n, y^n) \, p^{(\mu)}_{M_1|X_1^n}(m_1|x_1^n) p^{(\mu)}_{M_2|X_2^n}(m_2|x_2^n) p^{(\mu)}_{Y^n|M_1,M_2}(y^n|m_1, m_2)$$

$p^{(\mu)}_{M_j|X_j^n} : j \in [2], p^{(\mu)}_{Y^n|M_1,M_2}$ are the PMFs induced by the two randomized encoders and decoder respectively, corresponding to common randomness message μ. We let $\mathcal{R}_d(p_{\underline{X}Y})$ denote the set of achievable rate triples.

Our second main result is a characterization of $\mathcal{R}_d(p_{\underline{X}Y})$ which is provided in Theorem (2).

III. Soft Covering with Side Information

In this section, we provide a characterization of $\mathcal{R}_s(p_{XYZ})$.

Theorem 1. *$(R, C) \in \mathcal{R}_s(p_{XYZ})$ if and only if there exists a PMF p_{WXYZ} such that (i) $p_{XYZ}(x, y, z) = \sum_{w\in\mathcal{W}} p_{WXYZ}(w, x, y, z)$ for all (x, y, z) where $\mathcal{W}$ is the alphabet of W, (ii) $Z - X - W$ and $X - (Z, W) - Y$ are Markov chains, (iii) $|\mathcal{W}| \le (|\mathcal{X}||\mathcal{Y}||\mathcal{Z}|)^2$, and $R \ge I(X;W) - I(W;Z), R + C \ge I(XYZ;W) - I(W;Z)$.*

Proof. We provide the main elements (achievability in Sec. III-A and converse in Sec. IV) of our proof here with particular emphasis on the new elements. The reader is referred to [8] for more technical details.

A. Achievability

Throughout, $\mu \in [2^{nC}]$ denotes the C bits of common randomness shared between the encoder and decoder. For each $\mu \in [2^{nC}]$, we shall design a randomized encoder $E^{(\mu)} : \mathcal{X}^n \to [\Theta]$ and a randomized decoders $D^{(\mu)} : \mathcal{Z}^n \times [\Theta] \to \mathcal{Y}^n$ that induce PMFs $p^{(\mu)}_{M|X^n}$ and $p^{(\mu)}_{Y^n|Z^nM}$ respectively, for which

$$\mathscr{Q} := \frac{1}{2}\sum_{x^n,y^n,z^n} \left| P^n_{XYZ}(x^n, y^n, z^n) - \sum_{\mu\in[2^{nC}]} \sum_{m\in[\Theta]} \frac{p^n_{XZ}(x^n, z^n)}{2^{nC}} p^{(\mu)}_{M|X^n}(m|x^n) p^{(\mu)}_{Y^n|Z^n,M}(y^n|z^n, m) \right| \le \varepsilon. \quad (1)$$

The design of these randomized encoders and decoders involves building a codebook $\mathcal{C} = (\mathcal{C}^{(\mu)} : \mu \in [2^{nC}])$ where $\mathcal{C}^{(\mu)} = (\mathtt{w}^n(l, \mu) \in \mathcal{W}^n : l \in [2^{n\tilde{R}}])$ for $\mu \in [2^{nC}]$, $\mathcal{W}$ being the alphabet of W as in the theorem statement. Specifically, we let the codewords of $\mathcal{C}$ to be IID with distribution

$$\tilde{p}_{W^n}(w^n) = \frac{p^n_W(w^n)}{\sum_{w^n\in T_\delta(W)} p^n_W(w^n)} \text{ if } w^n \in T_\delta(W) \quad (2)$$

and 0 otherwise. On observing x^n, μ the randomized encoder chooses an index L in $[2^{n\tilde{R}}]$ according to a PMF $E^{(\mu)}_{L|X^n}(\cdot|\cdot)$. The chosen index is then mapped to an index in $[2^{nR}]$ which is communicated to the decoder. Before we specify the PMF $E^{(\mu)}_{L|X^n}(\cdot|\cdot)$, let us describe how the chosen index is mapped to an index in $[2^{nR}]$. In doing this, our first task is to identify and index the unique codewords in $\mathcal{C}$. Firstly, for $\mathcal{C}^{(\mu)}$, we let $\Theta^{(\mu)}$ denote the number of distinct codewords in $\mathcal{C}^{(\mu)}$. Secondly, we let $\mathcal{I}^{(\mu)}_{\mathcal{C}} : [2^{n\tilde{R}}] \to [\Theta^{(\mu)}]$ be defined such that $\mathcal{I}^{(\mu)}_{\mathcal{C}}(l) = \mathcal{I}^{(\mu)}_{\mathcal{C}}(\tilde{l})$ if and only if $\mathtt{w}(l, \mu) = \mathtt{w}(\tilde{l}, \mu)$. Lastly, we define a binning map $b^{(\mu)} : [\Theta^{(\mu)}] \to [2^{nR}]$. On observing x^n, the encoder chooses $L \in [2^{n\tilde{R}}]$ with respect to PMF $E^{(\mu)}_{L|X^n}(\cdot|x^n)$ and then communicates $b^{(\mu)}(\mathcal{I}^{(\mu)}_{\mathcal{C}}(L))$ to the decoder.

Before we specify $E^{(\mu)}_{L|X^n}(\cdot|\cdot)$ and characterize the induced PMF $p_{M|X^n}$ let us relate to the above three elements that make up the encoder. The PMF $E^{(\mu)}_{L|X^n}$ is analogous to the likelihood encoder $\Gamma_{J|X^n,K}$ of Cuff [3] but with important changes to incorporate typicality-based encoding that permits use of side-information at the decoder. The map $\mathcal{I}^{(\mu)}_{\mathcal{C}}$ eliminates duplication of indices with identical codewords and is employed for simplifying the analysis. The map $b^{(\mu)}$ performs standard information-theoretic binning [9] to utilize side-information.

We now specify $E^{(\mu)}_{L|X^n}(\cdot|\cdot)$. For $x^n \in T_\delta(X)$, let

$$E^{(\mu)}_{L|X^n}(l|x^n) = \frac{(1-\epsilon)p^n_{X|W}(x^n|\mathtt{w}^n(l,\mu))}{(1+\eta)2^{n\tilde{R}}p^n_X(x^n)} \quad \text{if } l \ne 0, \mathtt{w}^n(l,\mu) \in T_\delta(W|x^n), \text{ and} \quad (3)$$

$$E^{(\mu)}_{L|X^n}(l|x^n) = 1 - \sum_{l=1}^{2^{n\tilde{R}}} E^{(\mu)}_{L|X^n}(l|x^n) \text{ if } l = 0,$$

and $E^{(\mu)}_{L|X^n}(l|x^n) = 1_{\{l=0\}}$ for all $x^n \notin T_\delta(X)$. In specifying $E^{(\mu)}_{L|X^n}$, we have relaxed the requirement that $E^{(\mu)}_{L|X^n}(\cdot|x^n)$

be a PMF. This relaxation - a novelty of our work - yields analytical tractability of a random coding ensemble to be described in the sequel. However, note that these maps depend on the choice of the codebook $\mathcal{C}$. We prove in Appendix A that with high probability, $E^{(\mu)}_{L|X^n}(\cdot|x^n) : [2^{nC}] \to \mathbb{R}$ is a PMF for every $x^n \in \mathcal{X}^n$. This will form a part of our random codebook analysis and in fact, as we see in Lemma 2, one of the rate constraints is a consequence of the conditions necessary for the above definition of $E^{(\mu)}_{L|X^n}(\cdot|\cdot)$ to be a PMF. We also note that $E^{(\mu)}_{L|X^n}$ being a PMF guarantees $p_{M|X^n}$ is a PMF.

Having specified $E^{(\mu)}_{L|X^n}(\cdot|\cdot)$, we now characterize $p_{M|X^n}$.

$$p^{(\mu)}_{M|X^n}(m|x^n) = \sum_{w^n}\sum_{l=1}^{2^{n\tilde{R}}} E^{(\mu)}_{L|X^n}(w^n|x^n)\mathbb{1}_{\left\{\substack{w^n=\mathtt{w}^n(l,\mu)\\ b^{(\mu)}(\mathcal{I}^{(\mu)}_{\mathcal{C}}(l))=m}\right\}}$$
$$= \sum_{\substack{w^n\in\\ T_\delta(W|x^n)}} \sum_{l=1}^{2^{n\tilde{R}}} \frac{(1-\epsilon)p^n_{X|W}(x^n|w^n)}{2^{n\tilde{R}}(1+\eta)p^n_X(x^n)}\mathbb{1}_{\left\{\substack{w^n=\mathtt{w}^n(l,\mu)\\ b^{(\mu)}(\mathcal{I}^{(\mu)}_{\mathcal{C}}(l))=m}\right\}} \quad (4)$$

for $m \neq 0$ and $p^{(\mu)}_{M|X^n}(0|x^n) = 1 - \sum_{m=1}^{2^{nR}} p^{(\mu)}_{M|X^n}(m|x^n)$. We have thus described the encoder and $p_{M|X^n}$.

We now describe the decoder. On observing $z^n \in \mathcal{Z}^n, \mu$ and the index $m \in [2^{nR}]$ communicated by the encoder, the decoder populates $\mathcal{D}^{(\mu)}(z^n, m) = \{w^n \in \mathcal{W}^n : w^n = \mathtt{w}^n(l,\mu), b^{(\mu)}(\mathcal{I}^{(\mu)}_c(l)) = m, (w^n, z^n) \in T_\delta(W,Z)\}$. Let

$$f^{(\mu)}(m, z^n) = \begin{cases} w^n & \text{if } \mathcal{D}^{(\mu)}(z^n,m) = \{w^n\} \\ w_0 & \text{otherwise, i.e., } |\mathcal{D}^{(\mu)}(z^n,m)| \neq 1 \end{cases}.$$

The decoder chooses z^n according to PMF $p^n_{Y|ZW}(y^n|z^n, f(m,z^n))$. This implies we have the decoder given by

$$p^{(\mu)}_{Y^n|Z^nM}(\cdot|z^n, m) = p^n_{Y|WZ}(y^n|f^{(\mu)}(m,z^n), z^n). \quad (5)$$

We refer the reviewers to [8] for the remaining analysis.

IV. Converse

The proof for the converse follows from [1]. □

V. Distributed Soft Covering

Our main result is the following inner bound to $\mathcal{R}_d(p_{\underline{X}Y})$. In the following, we let $\underline{X} = (X_1, X_2), \underline{W} = (W_1, W_2), \underline{x} = (x_1, x_2)$ and $\underline{w} = (w_1, w_2)$.

Theorem 2. *Given a PMF $p_{X_1X_2Y}$, let $\mathcal{P}(p_{X_1X_2Y})$ denote the collection of all PMFs $p_{QW_1W_2\underline{X}Y}$ defined on $\mathcal{Q} \times \mathcal{W}_1 \times \mathcal{W}_2 \times \underline{\mathcal{X}} \times \mathcal{Y}$ such that (i) $p_{\underline{X}Y}(\underline{x}, y) = \sum_{(q,\underline{w})\in\mathcal{Q}\times\underline{\mathcal{W}}} p_{Q\underline{W}\underline{X}Y}(q, \underline{w}, \underline{x}, y)$ for all $(\underline{x}, y) \in (\underline{\mathcal{X}}, \mathcal{Y})$, (ii) $W_1 - Q X_1 - Q X_2 - W_2$ and $\underline{X} - Q\underline{W} - Y$ are Markov chains, (iii) $|\mathcal{W}_1| \leq |\mathcal{X}_1|$, $|\mathcal{W}_2| \leq |\mathcal{X}_2|$. Further, let $\beta(p_{Q\underline{W}\underline{X}Y})$ denote the set of rates and common randomness triple (R_1, R_2, C) that satisfy*

$$\begin{aligned} R_1 &\geq I(X_1; W_1|Q) - I(W_1; W_2|Q) \\ R_2 &\geq I(X_2; W_2|Q) - I(W_1; W_2|Q) \\ R_1 + R_2 &\geq I(X_1; W_1|Q) + I(X_2; W_2|Q) - I(W_1; W_2|Q) \\ R_1 + R_2 + C &\geq I(X_1X_2W_2Y; W_1|Q) + I(X_1X_2Y; W_2|Q) \\ &\quad - I(W_1; W_2|Q) \end{aligned} \quad (6)$$

where the mutual information terms are evaluated with the PMF $p_{QW_1W_2\underline{X}Y}$. We have

$$\textit{Closure}\left(\bigcup_{p_{Q\underline{W}\underline{X}Y}\in\mathcal{P}(p_{X_1X_2Y})} \beta(p_{Q\underline{W}\underline{X}Y})\right) \subseteq \mathcal{R}_d(P_{\underline{X}Y}). \quad (7)$$

In other words, (R_1, R_2, C) is achievable if $(R_1, R_2, C) \in \left(\bigcup_{p_{Q\underline{W}\underline{X}Y}\in\mathcal{P}(p_{X_1X_2Y})} \beta(p_{Q\underline{W}\underline{X}Y})\right)$.

In the interest of brevity, we only highlight the novelty in the design of random encoders and random decoder used. We refer the reader to [8] for a complete proof.

Proof. Having designed a randomized encoding scheme based on typicality, we are in a position to employ the same encoder for this distributed scenario. Let $\mu \in [2^{nC}]$ denote the common randomness shared amidst all terminals. The first encoder uses a part of the total common randomness available to it, say C_1 bits out of the C bits, denoted by $\mu_1 \in [2^{nC_1}]$. Similarly, let $\mu_2 \in [2^{nC_2}]$ denote the common randomness used by the second encoder. Our goal is to prove the existence of PMFs $p^{(\mu_1)}_{M_1|X_1^n}(m_1|x_1^n) : x_1^n \in \mathcal{X}_1^n, m_1 \in [\Theta_1], \mu_1 \in [2^{nC_1}], p^{\mu_2}_{M_2|X_2^n}(m_2|x_2^n) : x_2^n \in \mathcal{X}_2^n, m_2 \in [\Theta_2], \mu_2 \in [2^{nC_2}]$, and $p_{Y^n|M_1,M_2}(y^n|m_1,m_2) : y^n \in \mathcal{Y}^n, (m_1,m_2) \in [\Theta_1] \times [\Theta_2]$ such that

$$\begin{aligned} \mathscr{Q} := \frac{1}{2}\sum_{x_1^n,x_2^n,y^n} \Bigg| & p^n_{X_1X_2Y}(x_1^n, x_2^n, y^n) - \\ & \sum_{\mu\in[2^{nC}]}\sum_{m_1\in[\Theta_1]}\sum_{m_2\in[\Theta_2]} \frac{p^n_{X_1X_2}(x_1^n,x_2^n)}{2^{nC}} p^{(\mu_1)}_{M_1|X_1^n}(m_1|x_1^n) \\ & p^{(\mu_2)}_{M_2|X_2^n}(m_2|x_2^n) p^{(\mu)}_{Y^n|M_1,M_2}(y^n|m_1,m_2)\Bigg| \leq \varepsilon, \\ & \frac{\log\Theta_j}{n} \leq R_j + \epsilon : j \in [2] \end{aligned} \quad (8)$$

for sufficiently large n. Consider the collections $C_1 = (C_1^{(\mu_1)} : 1 \leq \mu_1 \leq 2^{nC_1})$ where $C_1^{(\mu_1)} = (\mathtt{w}_1(l_1,\mu_1) : 1 \leq l_1 \leq 2^{n\tilde{R}_1})$ and $C_2 = (C_2^{(\mu_2)} : 1 \leq \mu_2 \leq 2^{nC_2})$ where $C_2^{(\mu_2)} = (\mathtt{w}_2(l_2,\mu_2) : 1 \leq l_2 \leq 2^{n\tilde{R}_2})$. For this collection, we let

$$E^{(\mu_1)}_{L_1|X_1^n}(l_1|x_1^n) = \frac{1}{2^{n\tilde{R}_1}}\frac{1-\epsilon_1}{1+\eta}\sum_{w_1^n\in T_\delta(W_1|x_1^n)} \mathbb{1}_{\{\mathtt{w}^n(l_1,\mu_1)=w_1^n\}}\frac{p^n_{X_1|W_1}(x_1^n|w_1^n)}{p^n_{X_1}(x_1^n)}$$

$$E^{(\mu_2)}_{L_2|X_2^n}(l_2|x_2^n) = \frac{1}{2^{n\tilde{R}_2}}\frac{1-\epsilon_2}{1+\eta}\sum_{w_2^n\in T_\delta(W_2|x_2^n)} \mathbb{1}_{\{\mathtt{w}^n(l_2,\mu_2)=w_2^n\}}\frac{p^n_{X_2|W_2}(x_2^n|w_2^n)}{p^n_{X_2}(x_2^n)}$$

Further, we define maps $b_1^{(\mu_1)} : [2^{n\tilde{R}_1}] \to [2^{nR_1}]$ and $b_2^{(\mu_2)} : [2^{n\tilde{R}_2}] \to [2^{nR_2}]$ performing standard information-theoretic binning, with $0 < R_1 \leq \tilde{R}_1$ and $0 < R_2 \leq \tilde{R}_2$.

Using these maps, we induce the PMF $p^{(\mu_1)}_{M_1|X_1^n}$ on the message to be transmitted by the first encoder as

$$p^{(\mu_1)}_{M_1|X_1^n}(m_1|x_1^n) = \begin{cases} \mathbb{1}_{\{m_1=0\}} & \text{if } s_1^{(\mu_1)}(x_1^n) > 1, \\ 1 - s_1^{(\mu_1)}(x_1^n) & \text{if } m_1 = 0 \text{ and } s_1^{(\mu_1)}(x_1^n) \in [0,1], \\ \sum_{l_1=1}^{2^{n\tilde{R}_1}} E^{(\mu_1)}_{L_1|X_1^n}(l_1|x_1^n)\mathbb{1}_{\{b_1^{(\mu_1)}(l_1)=m_1\}} & \\ & \text{if } m_1 \neq 0 \text{ and } s_1^{(\mu_1)}(x_1^n) \in [0,1] \end{cases}$$

for all $x_1^n \in T_\delta(X_1)$ and $s_1^{(\mu_1)}(x_1^n)$ defined as $s_1^{(\mu_1)}(x_1^n) = \sum_{l_1=1}^{2^{n\tilde{R}_1}} E^{(\mu_1)}_{L_1|X_1^n}(l_1|x_1^n)$. For $x_1^n \notin T_\delta(X_1)$, we let $p^{(\mu_1)}_{M_1|X_1^n}(m_1|x_1^n) = \mathbb{1}_{\{m_1=0\}}$. Similarly, we define PMF $p^{(\mu_2)}_{M_2|X_2^n}$ for the second encoder as

$$p^{(\mu_2)}_{M_2|X_2^n}(m_2|x_2^n) = \begin{cases} \mathbb{1}_{\{m_2=0\}} & \text{if } s_2^{(\mu_2)}(x_2^n) > 1, \\ 1 - s_2^{(\mu_2)}(x_2^n) & \text{if } m_2 = 0 \text{ and } s_2^{(\mu_2)}(x_2^n) \in [0,1], \\ \sum_{l_2=1}^{2^{n\tilde{R}_2}} E^{(\mu_2)}_{L_2|X_2^n}(l_2|x_2^n)\mathbb{1}_{\{b_2(l_2)=m_2\}} & \\ & \text{if } m_2 \neq 0 \text{ and } s_2^{(\mu_2)}(x_2^n) \in [0,1]. \end{cases}$$

for all $x_2^n \in T_\delta(X_2)$ and $s_2^{(\mu_2)}(x_2^n)$ defined as $s_2^{(\mu_2)}(x_2^n) = \sum_{l_2=1}^{2^{n\tilde{R}_2}} E^{(\mu_2)}_{L_2|X_2^n}(l_2|x_2^n)$. For $x_2^n \notin T_\delta(X_2)$, we let $p^{(\mu_2)}_{M_2|X_2^n}(m_2|x_2^n) = \mathbb{1}_{\{m_2=0\}}$.

With this definition note that, $\sum_{m_1=0}^{2^{nR_1}} p^{(\mu_1)}_{M_1|X_1^n} = 1$ for all $\mu_1 \in [2^{nC_1}]$ and $x_1^n \in \mathcal{X}_1^n$ and similarly, $\sum_{m_2=0}^{2^{nR_2}} p^{(\mu_2)}_{M_2|X_2^n} = 1$ for all $\mu_2 \in [2^{nC_2}]$ and $x_2^n \in \mathcal{X}_2^n$.

We now describe the decoder. On observing μ and the indices $m_1, m_2 \in [2^{nR_1}] \times [2^{nR_2}]$ communicated by the encoder, the decoder first deduces (μ_1, μ_2) from μ and then populates $\mathcal{D}^{(\mu_1,\mu_2)}(m_1, m_2) = \{(w_1^n, w_2^n) \in \mathcal{W}_1^n \times \mathcal{W}_2^n : w_1^n = \mathtt{w}_1^n(l_1, \mu_1), w_2^n = \mathtt{w}_2^n(l_2, \mu_2), b_1^{(\mu_1)}(l_1) = m_1, b_2^{(\mu_2)}(l_2) = m_2, (w_1^n, w_2^n) \in T_\delta(W_1, W_2)\}$. Let

$$f^{(\mu)}(m_1, m_2) = \begin{cases} (w_1^n, w_2^n) & \text{if } \mathcal{D}^{(\mu_1,\mu_2)}(m_1, m_2) = \{(w_1^n, w_2^n)\} \\ (\tilde{w}_1^n, \tilde{w}_2^n) & \text{o.w. i.e., } |\mathcal{D}^{(\mu_1,\mu_2)}(m_1, m_2)| \neq 1 \end{cases}.$$

The decoder chooses y^n according to PMF $p^n_{Y|W_1W_2}(y^n|f^{(\mu)}(m_1, m_2))$. This implies the PMF $p^{(\mu_1)}_{Y^n|M_1M_2}(\cdot|\cdot)$ is given by

$$p^{(\mu)}_{Y^n|M_1M_2}(\cdot|m_1, m_2) = p^n_{Y|W_1W_2}(y^n|f^{(\mu)}(m_1, m_2)). \quad (9)$$

We now begin our analysis of (8). Our goal is to prove the existence of a collections c_1, c_2 for which (8) holds. We do this via random coding. Specifically, we prove that $\mathbb{E}[\mathscr{Q}] \leq \epsilon$ where the expectation is over the ensemble of codebooks. The PMF induced on the ensemble of codebooks is as specified below. The codewords of the random codebook $C_1^{(\mu_1)} = (\mathtt{W}_1(l_1, \mu_1) : 1 \leq l_1 \leq 2^{n\tilde{R}_1})$ for each $\mu_1 \in 2^{nC_1}$ are mutually independent and distributed with PMF

$$\mathbb{P}(\mathtt{W}_1(l_1, \mu_1) = w_1^n) = \frac{p^n_{W_1}(w_1^n)}{(1-\epsilon_1)}\mathbb{1}_{\{w_1^n \in T_\delta^n(W_1)\}}$$

Similarly, $C_2^{(\mu_2)} = (\mathtt{W}_2(l_2, \mu_2) : 1 \leq l_2 \leq 2^{n\tilde{R}_2})$ for each $\mu_2 \in [2^{nC_2}]$ are mutually independent and distributed as

$$\mathbb{P}(\mathtt{W}_2(l_2, \mu_2) = w_2^n) = \frac{p^n_{W_2}(w_2^n)}{(1-\epsilon_2)}\mathbb{1}_{\{w_2^n \in T_\delta^n(W_2)\}}$$

where $\epsilon_i = 1 - \mathbb{P}(T_\delta(W_i))$; $i = 1, 2$. We split $\mathscr{Q}$ into two terms using an indicator function $\mathbb{1}_{\{\mathtt{PMF}(C_1,C_2)\}}$ as

$$\begin{aligned} \mathbb{E}\mathscr{Q} &= \mathbb{E}\big[\mathscr{Q} \cdot \mathbb{1}_{\{\mathtt{PMF}(C_1,C_2)\}}\big] + \mathbb{E}\big[\mathscr{Q} \cdot \mathbb{1}^c_{\{\mathtt{PMF}(c_1,c_2)\}}\big] \\ &\leq \mathbb{E}\big[\mathscr{Q} \cdot \mathbb{1}_{\{\mathtt{PMF}(C_1,C_2)\}}\big] + 2 \cdot \mathbb{P}\left\{\mathbb{1}_{\{\mathtt{PMF}(C_1,C_2)\}} = 0\right\} \quad (10) \end{aligned}$$

where $\mathbb{1}_{\{\mathtt{PMF}(C_1,C_2)\}}$ is defined as

$$\mathbb{1}_{\{\mathtt{PMF}(C_1,C_2)\}} = \begin{cases} 1 & \text{if } s_1^{(\mu_1)}(x_1^n) \in [0,1] \text{ and } s_2^{(\mu_2)}(x_2^n) \in [0,1] \\ & \text{for all } x_1^n \in T_\delta(X_1), x_2^n \in T_\delta(X_2), \\ & \text{and for all } \mu_1 \in [2^{nC_1}], \mu_2 \in [2^{nC_2}] \\ 0 & \text{otherwise,} \end{cases}$$

and (10) follows from the upper bound of 1 over the total variation. Using the Lemma below, we prove that by constraining $\tilde{R}_1 \geq I(X_1; W_1) + 4\delta$ and $\tilde{R}_2 \geq I(X_2; W_2) + 4\delta$, $\mathbb{P}\left\{\mathbb{1}_{\{\mathtt{PMF}(C_1,C_2)\}} = 0\right\}$ can be made arbitrarily small. In other words, with high probability, we will have $E^{(\mu_1)}_{L_1|X_1^n}$ and $E^{(\mu_2)}_{L_2|X_2^n}$ such that $0 \leq \sum_{l_1=1}^{2^{n\tilde{R}_1}} E^{(\mu_1)}_{L_1|X_1^n} \leq 1$ for all $\mu_1 \in [2^{nC_1}]$ and $x_1^n \in T_\delta(X_1)$, and $0 \leq \sum_{l_2=1}^{2^{n\tilde{R}_2}} E^{(\mu_2)}_{L_2|X_2^n} \leq 1$ for all $\mu_2 \in [2^{nC_2}]$ and $x_2^n \in T_\delta(X_2)$,

Lemma 1. *For any $\delta, \eta \in (0, 1/2)$, if $\tilde{R}_1 > I(X_1 : W_1) + 4\delta_1$ and $\tilde{R}_2 > I(X_2 : W_2) + 4\delta_2$, where $\delta_1(\delta), \delta_2(\delta) \searrow 0$ as $\delta \searrow 0$, then*

$$\mathbb{P}\left[\left(\bigcap_{\mu=1}^{2^{nC_1}} \bigcap_{x^n \in T_\delta(X_1)} \left(E^{(\mu_1)}_{L_1|X_1^n}(l_1|x_1^n) \leq 1\right)\right) \bigcap \left(\bigcap_{\mu_2=1}^{2^{nC_2}} \bigcap_{x_2^n \in T_\delta(X_2)} \left(E^{(\mu_2)}_{L_2|X_2^n}(l_2|x_2^n) \leq 1\right)\right)\right] \to 1 \text{ as } n \to \infty$$

Proof. Using the lemma 2 from Appendix (A) twice, we get

$$\begin{aligned} &\mathbb{P}\left[\bigcap_{\mu_1=1}^{2^{nC}} \bigcap_{x_1^n \in T_{\delta(X_1)}} \left(\sum_{l_1=1}^{2^{n\tilde{R}_1}} E^{(\mu_1)}_{L_1|X_1^n}(l_1|x_1^n)\right) \leq 1\right] \\ &\geq 1 - 2 \cdot 2^{nC_1}|T_\delta(X_1)| \exp\left(-\frac{\eta^2 2^{n(\tilde{R}_1 - I(X_1,W_1) - 4\delta_1)}}{4\ln 2}\right), \\ &\mathbb{P}\left[\bigcap_{\mu=2}^{2^{nC_2}} \bigcap_{x_2^n \in T_{\delta(X_2)}} \left(\sum_{l_2=1}^{2^{n\tilde{R}_2}} E^{(\mu_2)}_{L_2|X_2^n}(l_2|x_2^n)\right) \leq 1\right] \\ &\geq 1 - 2 \cdot 2^{nC_2}|T_\delta(X_2)| \exp\left(-\frac{\eta^2 2^{n(\tilde{R}_2 - I(X_2,W_2) - 4\delta_2)}}{4\ln 2}\right). \end{aligned}$$

Applying union bound to the above inequalities gives,

$$\mathbb{P}\left[\left(\bigcap_{\mu=1}^{2^{nC}}\bigcap_{x^n\in T_\delta(X_1)}\left(E^{(\mu_1)}_{L_1|X_1^n}(l_1|x_1^n)\le 1\right)\right)\right.$$
$$\left.\bigcap\left(\bigcap_{\mu_2=1}^{2^{nC_2}}\bigcap_{x_2^n\in T_\delta(X_2)}\left(E^{(\mu_2)}_{L_2|X_2^n}(l_2|x_2^n)\le 1\right)\right)\right]$$
$$\ge 1-2\cdot 2^{nC_1}|T_\delta(X_1)|2\exp\left(-\frac{\eta^2 2^{n(\tilde{R}_1-I(X_1,W_1)-4\delta_1)}}{4\ln 2}\right)$$
$$-2\cdot 2^{nC_2}|T_\delta(X_2)|\exp\left(-\frac{\eta^2 2^{n(\tilde{R}_2-I(X_2,W_2)-4\delta_2)}}{4\ln 2}\right)$$

and hence the result follows. □

For the sake of brevity, we refer the readers to [8] where the complete details of the proof are provided. □

Appendix A
$E^{(\mu)}_{L|X^n}(\cdot|\cdot)$ is a PMF with high probability

Lemma 2. *For any $\delta,\eta\in(0,1)$, if $\tilde{R}>I(X:W)+4\delta$, then*

$$\mathbb{P}\left[\bigcap_{\mu=1}^{2^{nC}}\bigcap_{x^n\in T_\delta(X)}\left\{\sum_{l=1}^{2^{n\tilde{R}}}E^{(\mu)}_{L|X^n}(l|x^n)\le 1\right\}\right]\to 1 \text{ as } n\to\infty$$

Proof. From the definition of $E^{(\mu)}_{L|X^n}(l|x^n)$, we have for $x^n\in T_\delta(X)$, $\sum_{l=1}^{2^{n\tilde{R}}}E^{(\mu)}_{L|X^n}(l|x^n)=$

$$\frac{1}{2^{n\tilde{R}}}\left(\frac{1-\epsilon}{1+\eta}\right)\sum_{\substack{w^n\in\\ T_\delta(W|x^n)}}\sum_{l=1}^{2^{n\tilde{R}}}\mathbb{1}_{\{\mathtt{w}^n(l,\mu)=w^n\}}\frac{p^n_{X|W}(x^n|w^n)}{p^n_X(x^n)}.$$

Let us define $Z^{(\mu)}_l(x^n)$, for $x^n\in T_\delta(X)$ and $\mu\in[2^{nC}]$ as

$$Z^{(\mu)}_l(x^n)=(1-\epsilon)\sum_{w^n\in T_\delta(W|x^n)}\mathbb{1}_{\{\mathtt{w}^n(l,\mu)=w^n\}}p^n_{X|W}(x^n|w^n)$$

and let $D=2^{n(H(X|W)-\delta_1)}$, where $\delta_1(\delta)\searrow 0$ as $\delta\searrow 0$. This gives us the following bound on the expectation of the empirical average of $\{Z^{(\mu)}_l(x^n)\}_{l\in[2^{n\tilde{R}}]}$ as $\mathbb{E}\left[\frac{1}{N}\sum_{l=1}^N DZ^{(\mu)}_l(x^n)\right]$

$$=2^{n(H(X|W)-\delta)}\sum_{w^n\in T_\delta(W|x^n)}\tilde{p}^n_W(w^n)p^n_{X|W}(x^n|w^n)(1-\epsilon)$$
$$\ge 2^{n(H(X|W)-\delta)}2^{-n(H(X,W)+2\delta)}2^{n(H(W|X)-\delta)}$$
$$\ge 2^{-n(I(X,W)+4\delta)} \tag{11}$$

Further, we also have

$$DZ^{(\mu)}_l(x^n)\le 2^{n(H(X|W)-\delta)}2^{-n(H(X|W)-\delta)}(1-\epsilon)\sum_{w^n\in T_\delta(W|x^n)}\mathbb{1}_{\{W^n(l,\mu)=w^n\}}\le 1 \tag{12}$$

Since, for every $x^n\in T_\delta(X)$ and $\mu\in[2^{nC}]$, we have $\{Z^{(\mu)}_l(x^n)\}_l$ to be a sequence of IID Random variables, we can approximate its empirical average, using a more refined Chernoff-Hoeffding bound given by

Lemma 3. *Let $\{Z_n\}_{n=1}^N$ be a sequence of N IID random variables bounded as $Z_n\in[0,1]\ \ \forall n\in[N]$, and suppose $\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^N Z_n\right]=\mu$ be lower bounded by a positive constant θ as $\mu\ge\theta$ where $\theta\in(0,1)$, then for every $\eta\in(0,1/2)$ and $(1+\eta)\theta<1$, we can bound the probability that the ensemble average of the sequence $\{Z_n\}_{n=1}^N$ lies in $(1\pm\eta)\mu$ as*

$$\mathbb{P}\left(\frac{1}{N}\sum_{n=1}^N Z_n\in[(1-\eta)\mu,(1+\eta)\mu]\right)\ge 1-2\exp\left(-\frac{N\eta^2\theta}{4\ln 2}\right)$$

Proof. Follows from Operator Chernoff Bound [10]. □

Note that, from (11 and 12), $\{DZ^{(\mu)}_l(x^n)\}_l$ satisfies the constraints needed in the above lemma . Thus applying Lemma (3) to $\{DZ^{(\mu)}_l(x^n)\}_l$ gives

$$\mathbb{P}\left(\frac{1}{N}\sum_{l=1}^N Z^{(\mu)}_l(x^n)\in\left[(1-\eta)\mathbb{E}\left[Z(x^n)\right],(1+\eta)\mathbb{E}\left[Z(x^n)\right]\right]\right)$$
$$\ge 1-2\exp\left(-\frac{\eta^2 2^{n(R-I(X,W)-4\delta)}}{4\ln 2}\right) \tag{13}$$

where $Z^{(\mu)}(x^n)$ denotes the generic random variable from the IID sequence $\{Z^{(\mu)}_l(x^n)\}_l$. Substituting the following simplification

$$\frac{1}{2^{n\tilde{R}}}\sum_{l=1}^{2^{n\tilde{R}}}Z^{(\mu)}_l(x^n)=(1+\eta)p^n_X(x^n)\sum_{l=1}^{2^{n\tilde{R}}}E^{(\mu)}_{L|X^n}(l|x^n)$$

which follows from the definition of $Z^n_l(x^n)$, into (13) gives

$$\mathbb{P}\left((1+\eta)p^n_X(x^n)\sum_{l=1}^{2^{n\tilde{R}}}E^{(\mu)}_{L|X^n}(l|x^n)\le(1+\eta)\mathbb{E}[Z^{(\mu)}(x^n)]\right)$$
$$\ge 1-2\exp\left(-\frac{\eta^2 2^{n(\tilde{R}-I(X,W)-4\delta_1)}}{4\ln 2}\right)$$

Further we can bound $\mathbb{E}[Z^{(\mu)}(x^n)]$ as

$$\frac{\mathbb{E}[Z^{(\mu)}(x^n)]}{p^n_X(x^n)}\le\frac{(1-\epsilon)}{p^n_X(x^n)}\sum_{w^n\in T_\delta(W|x^n)}\tilde{p}^n_W(w^n)p^n_{X|W}(x^n|w^n)\le 1$$

where the last inequality above is obtained by adding more terms in the summation This simplifies as

$$\mathbb{P}\left(\sum_{l=1}^{2^{n\tilde{R}}}E^{(\mu)}_{L|X^n}(l|x^n)\le 1\right)\ge 1-2\exp\left(-\frac{\eta^2 2^{n(\tilde{R}-I(X,W)-4\delta_1)}}{4\ln 2}\right)$$

Using union bound, for all $\mu\in[2^{nC}]$ and $x^n\in T_\delta(X)$ gives,

$$\mathbb{P}\left[\bigcap_{\mu=1}^{2^{nC}}\bigcap_{x^n\in T_{\delta(X)}}\left(\sum_{l=1}^{2^{n\tilde{R}}}E^{(\mu)}_{L|X^n}(l|x^n)\right)\le 1\right]$$
$$\ge 1-2\cdot 2^{nC}|T_\delta(X)|\exp\left(-\frac{\eta^2 2^{n(\tilde{R}-I(X,W)-4\delta_1)}}{4\ln 2}\right)$$

and hence the proof completes. □

REFERENCES

[1] M. H. Yassaee, A. Gohari, and M. R. Aref, "Channel simulation via interactive communications," IEEE Transactions on Information Theory, vol. 61, no. 6, pp. 2964–2982, 2015.

[2] A. Wyner, "The common information of two dependent random variables," IEEE Transactions on Information Theory, vol. 21, no. 2, pp. 163–179, March 1975.

[3] P. Cuff, "Distributed channel synthesis," IEEE Transactions on Information Theory, vol. 59, no. 11, pp. 7071–7096, Nov 2013.

[4] P. W. Cuff, "Communication in networks for coordinating behavior," Ph.D. dissertation, Stanford, CA, USA, 2009.

[5] A. Winter, ""extrinsic"and "intrinsic"data in quantum measurements: Asymptotic convex decomposition of positive operator valued measures," Communications in mathematical physics, vol. 244, no. 1, pp. 157–185, 2004.

[6] M. M. Wilde, P. Hayden, F. Buscemi, and M.-H. Hsieh, "The information-theoretic costs of simulating quantum measurements," Journal of Physics A: Mathematical and Theoretical, vol. 45, no. 45, p. 453001, 2012.

[7] M. Heidari, T. A. Atif, and S. Sandeep Pradhan, "Faithful simulation of distributed quantum measurements with applications in distributed rate-distortion theory," in 2019 IEEE International Symposium on Information Theory (ISIT), July 2019, pp. 1162–1166.

[8] T. A. Atif, A. Padakandla, and S. S. Pradhan, "Source coding for synthesizing correlated randomness," arXiv preprint arXiv:2004.03651, 2020.

[9] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," vol. 22, no. 1, pp. 1–10, January 1976.

[10] R. Ahlswede and A. Winter, "Strong converse for identification via quantum channels," IEEE Transactions on Information Theory, vol. 48, no. 3, pp. 569–579, 2002.