A Technique for Deriving One-Shot Achievability Results in Network Information Theory

Mohammad Hossein Yassaee, Mohammad Reza Aref, Amin Gohari

Information Systems and Security Lab (ISSL),

Sharif University of Technology, Tehran, Iran,

E-mail: yassaee@ee.sharif.edu, {aref,aminzadeh}@sharif.edu.

Abstract

This paper proposes a novel technique to prove a one-shot version of achievability results in network information theory. The technique is not based on covering and packing lemmas. In this technique, we use an stochastic encoder and decoder with a particular structure for coding that resembles both the ML and the joint-typicality coders. Although stochastic encoders and decoders do not usually enhance the capacity region, their use simplifies the analysis. The Jensen inequality lies at the heart of error analysis, which enables us to deal with the expectation of many terms coming from stochastic encoders and decoders at once. The technique is illustrated via several examples: point-to-point channel coding, Gelfand-Pinsker, Broadcast channel (Marton), Berger-Tung, Heegard-Berger/Kaspi, Multiple description coding and Joint source-channel coding over a MAC. Most of our one-shot results are new. The asymptotic forms of these expressions is the same as that of classical results. Our one-shot bounds in conjunction with multi-dimensional Berry-Essen CLT imply new results in the finite blocklength regime. In particular applying the one-shot result for the memoryless broadcast channel in the asymptotic case, we get the entire region of Marton's inner bound without any need for time-sharing.

I. INTRODUCTION

Information theory aims to find optimal reliable communication rates in networks. The combinatorial structure of networks makes the problem difficult in general. However one can employ law of large numbers by looking at asymptotic behavior of networks for large blocklengths. But this comes at the cost of a long delay. This motivates looking at the problem in the so called "finite blocklength regime." The blocklength in this regime is not infinitely long, but is sufficiently large for certain CLTs to hold. Originally studied by Strassen [1], there has been a recent surge of works on this topic following the results of Polyanskiy et al [2] (see for instance [3]-[6]).

In this paper we consider *one-shot* network information theory where a *single* use of the network is allowed. In this case the probability of error cannot necessarily be driven to zero. Further, well-known techniques such as joint typicality and time sharing are not applicable here. Given an admissible probability of error, our goal is to find a characterization of a set of achievable rates that resembles the form of the asymptotic results. There has been some previous work along this direction. Wang and Renner [7] derive one-shot upper and lower bounds for the problem of transmission of classical information over a classic-quantum channel (see also [8]). Recently Verdu has proposed a one-shot version of the covering and packing lemmas, and has applied it to a set of classical problems in information theory [9].

Our main contribution is a proof technique for deriving the results on the one-shot region. The technique uses elementary tools and is not based on extensions of packing or covering lemmas. It is based on a particular construction for encoder and decoders that is not ML, but resembles both the ML and jointly typical coders. Our proposed decoders are stochastic and intuitively attempt to pass the received symbol through a certain inverse conditional distribution. The Jensen's inequality is central to the analysis of the error. The technique can be widely applied to problems of network information theory. To illustrate this, we derive new results for the problems of Gelfand-Pinsker, broadcast channel, joint-source channel coding over MAC, Berger-Tung, Heegard-Berger/Kaspi and multiple description coding. The asymptotic forms of these expressions is the same as that of classical results. Our one-shot bounds also imply new results in the finite blocklength regime.

The most related previous work is that of Verdu [9]. Whereas [9] proposes a one-shot covering and packing lemmas to solve network problems, we propose a direct analysis comprising of a chain of inequalities. By bypassing the need for covering and packing lemmas, we can provide bounds for problems that were originally solved using mutual covering and packing lemmas in the asymptotic regime. This is helpful because no one-shot extension of the mutual covering and packing lemma exists. We discuss this point in more details in Remark 3.

This paper is organized as follows: In Section II we provide some definitions. This is followed by three sections that provide application of the technique to different scenarios. In Section III we consider three problems of multi-terminal channel coding, namely, point-to-point channel, Gelfand-Pinsker and broadcast channel (Marton). In Section IV we consider three problems of lossy multi-terminal source coding, i.e. Berger-Tung, Heegard-Berger/Kaspi and Multiple description coding. Lastly in Section V we study a joint source-channel coding problem of transmission correlated sources over a MAC. In each of these problems we provide a one-shot achievability result. Corresponding finite blocklength results could be derived from these results. To illustrate this, we have derived such bounds for the Gelfand-Pinsker and broadcast channel problems.

II. DEFINITIONS

Definition 1. *Given a pmf $p_{X,Y}$, the conditional information of x given y is defined by*

$$h_p(x|y) = \log \frac{1}{p_{X|Y}(x|y)}.$$

Definition 2. *For a pmf $p_{X,Y,Z}$, the* conditional information density $i(x; y|z)$ *is defined by*

$$i_p(x; y|z) := \log \frac{p(x, y|z)}{p(x|z)p(y|z)},$$

and for general r.v.'s it is defined by

$$i_p(x; y|z) := \log \frac{\mathsf{d}p_{X,Y|Z}}{\mathsf{d}(p_{X|Z} \times p_{Y|Z})}(x, y, z).$$

Whenever the underlying distribution is clear from the context, we drop the subscript p from $i_p(x; y|z)$.

Definition 3. *Let* **X** *be a multi-dimensional normal variable with zero mean and covariance matrix* V. *The complementary multivariate Gaussian cumulative distribution region associated with* V *is defined by*

$$\mathcal{Q}^{-1}(\mathsf{V}, \epsilon) := \{\mathbf{x} : \mathsf{P}(\mathbf{X} \leq \mathbf{x}) \geq 1 - \epsilon\}.$$

We use M and J to denote size of alphabets of random variables M and J, respectively, i.e. $\mathsf{M} = |\mathcal{M}|$ and $\mathsf{J} = |\mathcal{J}|$. All the logarithms are in base two throughout this paper.

III. MULTI-TERMINAL CHANNEL CODING PROBLEMS

To illustrate the application of our technique to multi-terminal channel coding problems, we study the problems of point-to-point channel, Gelfand-Pinsker and Broadcast channels (Marton) in this section.

A. Point-to-point channel

We begin our illustration of the one-shot achievability proof with the classical point-to-point channel. Consider a channel with the law $q_{Y|X}$ and an input distribution q_X. Let $\mathcal{C} = \{X(1), \cdots, X(\mathsf{M})\}$ be a random codebook where the elements $X(i)$ are drawn independently from q_X (each codeword $X(i)$ is only a single rv). As usual, $X(m)$ is the codeword used for transmission of the message m. For the decoding we use an stochastic variation of MAP decoding. Instead of declaring the message $\hat{m}$ with maximal posterior probability as in MAP, the decoder randomly draws a message $\hat{m}$ from the conditional pmf $P_{M|Y}$, where P is the induced pmf by the code, $P_{M,Y}(m, y) = \frac{1}{\mathsf{M}}q(y|X(m))$.[1] More specifically,

$$P_{M|Y}(\hat{m}|y) = \frac{q(y|X(\hat{m}))}{\sum_{\bar{m}} q(y|X(\bar{m}))} = \frac{2^{i_q(y;X(\hat{m}))}}{\sum_{\bar{m}} 2^{i_q(y;X(\bar{m}))}}. \tag{1}$$

The mutual information term $i_q(y; X(\hat{m}))$ is computed using pmf $q_X q_{Y|X}$ that has nothing to do with the pmf induced by the code. However the sequence $X(\hat{m})$ itself is random, hence we have used $P_{M|Y}(\hat{m}|y)$ (capital P) to denote the pmf.

We refer this decoder as *stochastic likelihood coder* (SLC), or as *stochastic mutual information coder* (SMC).[2] The second equality shows that the probability of selecting a message is proportional to two to the power of its

---

[1]The pmf is random due to the random codebook.

[2]The reason for introducing two names for apparently the same object will become clear later. These decoders will not be the same in other problems.

mutual information with the received output. So codewords with higher mutual information have a higher chance of being selected as the output of the decoder. This resembles the widely used joint typicality decoder in the asymptotic regime.

Theorem 1. *The* expected value *of the probability of* correct decoding *of SLC (or SMC) for a randomly generated codebook of size* M *is bounded from below by*

$$\mathbb{E}_{\mathcal{C}} \mathsf{P}[C] \geq \mathbb{E}_{q_{XY}} \frac{1}{1 + (\mathsf{M} - 1)2^{-\imath_q(X;Y)}}. \tag{2}$$

Proof: Observe that the joint distribution of random variables factors as,

$$P_{MY\hat{M}}(m, y, \hat{m}) = \frac{1}{\mathsf{M}} q(y|X(m)) P_{M|Y}(\hat{m}|y),$$

and the probability of correct decoding can be written as $\mathsf{P}[C] = \sum_{m,y} P_{MY\hat{M}}(m, y, m)$, hence we have:

$$\mathbb{E}\mathsf{P}[C] = \mathbb{E} \sum_{m,y} \frac{1}{\mathsf{M}} q(y|X(m)) \frac{2^{\imath_q(y;X(m))}}{\sum_{\bar{m}} 2^{\imath_q(y;X(\bar{m}))}} \tag{3}$$

$$= \mathbb{E} \sum_{y} q(y|X(1)) \frac{2^{\imath_q(y;X(1))}}{\sum_{\bar{m}} 2^{\imath_q(y;X(\bar{m}))}} \tag{4}$$

$$= \sum_{y} \mathbb{E}_{X(1)} \mathbb{E}_{\mathcal{C}|X(1)} q(y|X(1)) \frac{2^{\imath_q(y;X(1))}}{\sum_{\bar{m}} 2^{\imath_q(y;X(\bar{m}))}} \tag{5}$$

$$\geq \sum_{y} \mathbb{E}_{X(1)} q(y|X(1)) \frac{2^{\imath_q(y;X(1))}}{\mathbb{E}_{\mathcal{C}|X(1)} \sum_{\bar{m}} 2^{\imath_q(y;X(\bar{m}))}} \tag{6}$$

$$= \sum_{y} \mathbb{E}_{X(1)} q(y|X(1)) \frac{2^{\imath_q(y;X(1))}}{2^{\imath_q(y;X(1))} + (\mathsf{M} - 1)} \tag{7}$$

$$= \sum_{x,y} q(x) q(y|x) \frac{2^{\imath_q(y;x)}}{2^{\imath_q(y;x)} + (\mathsf{M} - 1)} \tag{8}$$

$$= \mathbb{E}_{XY} \frac{1}{1 + (\mathsf{M} - 1)2^{-\imath_q(X;Y)}}, \tag{9}$$

where (5) follows from the rule of iterated expectation, (6) follows from the Jensen inequality for the convex function $f(x) = \frac{1}{x}$ on $\mathbb{R}_+$, and (7) follows from the following equation for $\bar{m} \neq 1$,

$$\mathbb{E}_{\mathcal{C}|X(1)} 2^{\imath(y;X(\bar{m}))} = \sum_{x} q(x) 2^{\imath(y;x)} = \sum_{x} q(x|y) = 1,$$

where we use the fact that $X(\bar{m})$ is independent of $X(1)$ for $\bar{m} \neq 1$ and is drawn from q_X. ∎

B. Gelfand-Pinsker

Consider the problem of transmitting a message over a state-dependent channel with state information available at the encoder, depicted in Fig. 1. Let q_S and $q_{Y|X,S}$ be the state's pmf and the channel transition probability, respectively.
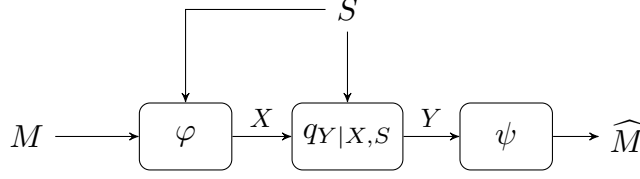
Fig. 1. Coding over a state-dependent channel.

1) One-shot achievability:

Definition 4. *An* M-*code for state-dependent channel consists of a (possibly stochastic) encoder* $\varphi : [1 : \mathsf{M}] \times \mathcal{S} \mapsto \mathcal{X}$, *and a (possibly stochastic) decoder* $\psi : \mathcal{Y} \mapsto [1 : \mathsf{M}]$.

Theorem 2. *Given any* $q_{U|S}$ *and function* $x(u, s)$, *there is an* M-*code for a single use of the channel whose probability of correct decoding is bounded from below by*

$$\mathbb{E}_{USY} \frac{1}{(1 + \mathsf{J}^{-1} 2^{i_q(U;S)})(1 + \mathsf{MJ} 2^{-i_q(U;Y)})}, \tag{10}$$

where $\mathsf{J} > 0$ *is an arbitrary integer and* $q(u, s, x, y) = q(s)q(u|s)\mathbf{1}[x = x(u, s)]q(y|x, s)$. *Moreover, loosening this bound gives the following upper bound on the error probability of the code,*

$$\mathsf{P}[\log \mathsf{J} - i_q(U; S) < \gamma , \text{ or}$$

$$i_q(U; Y) - \log \mathsf{MJ} < \gamma] + 3 \times 2^{-\gamma}, \tag{11}$$

where γ *is any positive number.*

Remark 1. *If we apply the above result to* n *copies of a memoryless state dependent channel, we recover the asymptotic Gelfand-Pinsker result. In this derivation the first term in the denominator of* (10), $1 + \mathsf{J}^{-1} 2^{i(U;S)}$ *corresponds to a covering lemma in the asymptotic case, while the second term* $1 + \mathsf{MJ} 2^{-i(U;Y)}$ *corresponds to a packing lemma. Observe that the first term is proportional to* J^{-1} *whereas the second term is proportional to* MJ. *Thus the above formula combines covering and packing lemmas at once.*

Remark 2. *If we further loosen the first term of eq.* (11) *using the union bound, we get Verdu's bound on this problem [9] except for the term* $3 \times 2^{-\gamma}$. *This residual term is not of significance in the finite blocklength* n-*letter regime where we choose* γ *of the order of* $\log(n)$ *(see Theorem 3); the main contribution comes from the probability terms. In a concurrent work [10], Watanabe et.al., prove an expression similar to eq.* (11) *using a different method based on channel simulation. They also applied their approach to the problem of source coding with a helper and to the Wyner-Ziv problem. It is not clear whether their approach is applicable to the scenarios such as broadcast channel, multiple description coding, etc that is solved in the asymptotic case using the multivariate covering lemma, since no extension of channel simulation (based on the work of Cuff [11]) is known for multiuser scenarios.*

Nonetheless, our technique bypasses the need for either an extension of covering lemma to multivariate covering, or a multi-terminal extension of the channel simulation result. See also Remark 3.

Proof:

Let $\mathcal{C} = \{U(m,j)\}_{m=1,j=1}^{\mathsf{M},\mathsf{J}}$ be a random codebook whose elements are drawn independently from q_U. Here J is introducing redundancy but since it will be decoded at the receiver we can view it as a dummy message.

Encoding: Instead of using conventional random covering, we use an SMC which acts as follows. Given m and s, the SMC chooses an index j with the probability

$$P_{\mathsf{Enc}}(j|m,s) = \frac{2^{\imath_q(s;U(m,j))}}{\sum_{\tilde{j}} 2^{\imath_q(s;U(m,\tilde{j}))}}.$$

Then the encoder transmits $x(U(m,j),s)$ through the channel. Observe that the above SMC resembles a joint-typical encoder of the asymptotic regime. Given m and s, the higher the information between $U(m,j)$ and s, the more likely we choose it at the encoder.

Decoding: In contrast to the point-to-point problem, computing the error probability of SLC is challenging. An SLC uses the induced $P_{M,J|Y}$ by the code. Instead, we use an SMC with the following rule for decoding. Observing y, decoder uses the following SMC to find both the message m and the dummy message j:

$$P_{\mathsf{Dec}}(\hat{m},\hat{j}|y) = \frac{2^{\imath_q(y;U(\hat{m},\hat{j}))}}{\sum_{\bar{m},\bar{j}} 2^{\imath_q(s;U(\bar{m},\bar{j}))}}.$$

Analysis: We declare an error if $(\hat{m},\hat{j}) \neq (m,j)$. Observe that the joint distribution of random variables factors as,

$$P_{MJSY\hat{M}\hat{J}}(m,j,s,y,\hat{m},\hat{j}) = \frac{1}{\mathsf{M}} q(s) P_{\mathsf{Enc}}(j|m,s) q(y|U(m,j),s) P_{\mathsf{Dec}}(\hat{m},\hat{j}|y),$$

where $q(y|U(m,j),s) = q_{Y|X,S}\big(y|x(U(m,j),s),s\big)$. The probability of correct decoding is $\mathsf{P}[C] = \sum_{m,j,s,y} P_{MJSY\hat{M}\hat{J}}(m,j,s,y,m,j)$; hence we have:

$$\mathbb{E}\mathsf{P}[C] = \mathbb{E} \sum_{m,j,s,y} \frac{1}{\mathsf{M}} q(s) \frac{2^{\imath(s;U(m,j))}}{\sum_{\tilde{j}} 2^{\imath(s;U(m,\tilde{j}))}} q(y|U(m,j),s) \frac{2^{\imath(y;U(m,j))}}{\sum_{\bar{m},\bar{j}} 2^{\imath(y;U(\bar{m},\bar{j}))}} \tag{12}$$

$$= \mathbb{E} \sum_{s,y} \mathsf{J} q(s) \frac{2^{\imath(s;U(1,1))}}{\sum_{\tilde{j}} 2^{\imath(s;U(1,\tilde{j}))}} q(y|U(1,1),s) \frac{2^{\imath(y;U(1,1))}}{\sum_{\bar{m},\bar{j}} 2^{\imath(y;U(\bar{m},\bar{j}))}} \tag{13}$$

$$\geq \sum_{s,y} \mathbb{E}_{U(1,1)} \left(\frac{\mathsf{J} q(s) 2^{\imath(s;U(1,1))}}{\mathbb{E}_{\mathcal{C}|U(1,1)} \sum_{\tilde{j}} 2^{\imath(s;U(1,\tilde{j}))}} q(y|U(1,1),s) \frac{2^{\imath(y;U(1,1))}}{\mathbb{E}_{\mathcal{C}|U(1,1)} \sum_{\bar{m},\bar{j}} 2^{\imath(y;U(\bar{m},\bar{j}))}} \right) \tag{14}$$

$$\geq \sum_{s,y} \mathbb{E}_{U(1,1)} \left(\frac{\mathsf{J} q(s) 2^{\imath(s;U(1,1))}}{2^{\imath(s;U(1,1))} + \mathsf{J}} q(y|U(1,1),s) \frac{2^{\imath(y;U(1,1))}}{2^{\imath(y;U(1,1))} + \mathsf{MJ}} \right) \tag{15}$$

$$= \sum_{u,s,y} q(u,s,y) \frac{\mathsf{J}}{2^{\imath(s;u)} + \mathsf{J}} \cdot \frac{2^{\imath(y;u)}}{2^{\imath(y;u)} + \mathsf{MJ}} \tag{16}$$

$$= \mathbb{E}_{USY} \frac{1}{(1 + \mathsf{J}^{-1} 2^{\imath(U;S)})(1 + \mathsf{MJ} 2^{-\imath(U;Y)})}, \tag{17}$$

where (13) is due to symmetry, the <u>main step</u> (14) follows from Jensen inequality for the two-valued convex function $f(x_1, x_2) = \dfrac{1}{x_1 x_2}$ on $\mathbb{R}_+^2$, (15) follows from the fact that $U(i,j)$ is independent of $U(1,1)$ for $(i,j) \neq (1,1)$ and generated according to q_U, and (16) follows from the fact that $U(1,1)$ is generated according to q_U.

Deriving the loosened bound (11):

$$\mathbb{E}_{USY} \frac{1}{(1 + \mathsf{J}^{-1}2^{\imath(U;S)})(1 + \mathsf{MJ}2^{-\imath(U;Y)})} \geq \mathbb{E}_{USY} \frac{\mathbf{1}\{\log \mathsf{J} - \imath(U;S) \geq \gamma, \ \imath(U;Y) - \log \mathsf{MJ} \geq \gamma\}}{(1 + \mathsf{J}^{-1}2^{\imath(U;S)})(1 + \mathsf{MJ}2^{-\imath(U;Y)})} \tag{18}$$

$$\geq \frac{1}{(1 + 2^{-\gamma})^2} \mathsf{P}\left[\log \mathsf{J} - \imath(U;S) \geq \gamma, \ \imath(U;Y) - \log \mathsf{MJ} \geq \gamma\right] \tag{19}$$

$$\mathsf{P}[\mathcal{E}] \leq 1 - \frac{1}{(1 + 2^{-\gamma})^2} \mathsf{P}\left[\log \mathsf{J} - \imath(U;S) \geq \gamma, \ \imath(U;Y) - \log \mathsf{MJ} \geq \gamma\right] \tag{20}$$

$$= \mathsf{P}\left[\log \mathsf{J} - \imath(U;S) < \gamma, \ \text{or} \ \imath(U;Y) - \log \mathsf{MJ} < \gamma\right]$$

$$+ \left(1 - \frac{1}{(1 + 2^{-\gamma})^2}\right) \mathsf{P}\left[\log \mathsf{J} - \imath(U;S) \geq \gamma, \ \imath(U;Y) - \log \mathsf{MJ} \geq \gamma\right] \tag{21}$$

$$\leq \mathsf{P}\left[\log \mathsf{J} - \imath(U;S) < \gamma, \ \text{or} \ \imath(U;Y) - \log \mathsf{MJ} < \gamma\right] + \left(1 - \frac{1}{(1 + 2^{-\gamma})^2}\right) \tag{22}$$

$$\leq \mathsf{P}\left[\log \mathsf{J} - \imath(U;S) < \gamma, \ \text{or} \ \imath(U;Y) - \log \mathsf{MJ} < \gamma\right] + 3 \times 2^{-\gamma}. \tag{23}$$

∎

2) Second Order achievability of Gelfand-Pinsker channel:

Theorem 3. *Given a memoryless state-dependent channel $(q_S, q_{Y|X,S})$ with state known non-causally at the encoder, for any $(q_{U|S}, x(u,s))$, the following rate is (n, ϵ)-achievable:*

$$R = I(U;Y) - I(U;S) - \frac{1}{\sqrt{n}} R_D - O\left(\frac{\log n}{n}\right) \tag{24}$$

where

$$R_D = \min_{R : \exists \tilde{R}, \ \text{s.t.} \ [\tilde{R}, R - \tilde{R}]^\mathsf{T} \in \mathcal{Q}^{-1}(\mathbb{V}_{\mathsf{GP}}, \epsilon)} R, \tag{25}$$

and

$$\mathbb{V}_{\mathsf{GP}} = \mathsf{Cov}\begin{bmatrix} \imath(U;S) \\ \imath(U;Y) \end{bmatrix}. \tag{26}$$

Proof: We apply (11) to n use of the channel. Assume that $q_{U^n}(u^n) = \prod_{i=1}^n q_U(u_i)$, so (U^n, S^n, Y^n) are i.i.d.. Substituting $\gamma = \frac{1}{2}\log n$ in (11) implies:

$$P[\mathcal{E}] \leq \mathsf{P}\left(\log \mathsf{J} - \imath(U^n; S^n) \leq \frac{1}{2}\log n, \ \text{or} \ \imath(U^n; Y^n) - \log \mathsf{MJ} \leq \frac{1}{2}\log n\right) + \frac{3}{\sqrt{n}}. \tag{27}$$

Given $\epsilon > 0$, finding a code such that for some $\epsilon' \leq \epsilon$

$$\epsilon' = \mathsf{P}\left(\log \mathsf{J} - \imath(U^n; S^n) \leq \frac{1}{2}\log n, \ \text{or} \ \imath(U^n; Y^n) - \log \mathsf{MJ} \leq \frac{1}{2}\log n\right) + \frac{3}{\sqrt{n}}. \tag{28}$$
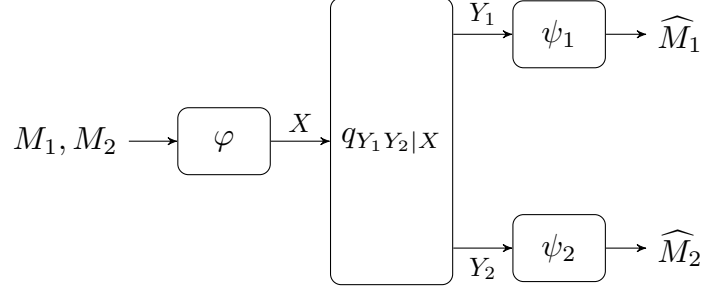
Fig. 2.   Coding over a broadcast channel.

implies that ϵ is an upper bound on $P[\mathcal{E}]$. Equivalently, we would like to find a code such that

$$1 - \epsilon' - \frac{3}{\sqrt{n}} = \mathsf{P}\left(\log \mathsf{J} - \imath(U^n; S^n) > \frac{1}{2}\log n, \quad \imath(U^n; Y^n) - \log \mathsf{MJ} > \frac{1}{2}\log n\right). \tag{29}$$

Let $\log \mathsf{J} = nI(U; S) + \sqrt{n}\tilde{R} + \frac{1}{2}\log n$ and $\log \mathsf{M} = n(I(U; Y) - I(U; S)) - \sqrt{n}R - \log n$. The random variables $\imath(U^n; S^n)$ and $\imath(U^n; Y^n)$ are sum of i.i.d. random variabels. Applying multi-dimensional Berry-Essen CLT [20] to (29) implies the following equivalent form:

$$1 - \epsilon' - O(\frac{1}{\sqrt{n}}) = \mathsf{P}_G\left(\begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \leq \begin{bmatrix} \tilde{R} \\ R - \tilde{R} \end{bmatrix}\right), \tag{30}$$

where $G = [G_1\ G_2]^\mathsf{T}$ is a multidimensional normal r.v. with zero mean and $\mathsf{Cov}G = \mathbb{V}_{\mathsf{GP}}$. Using the definition of $\mathcal{Q}^{-1}(\mathsf{V}, \epsilon)$ and smoothness of distribution of normal r.v., we get

$$[\tilde{R}, R - \tilde{R}]^\mathsf{T} \in \mathcal{Q}^{-1}(\mathbb{V}_{\mathsf{GP}}, \epsilon') + O\left(\frac{\log n}{\sqrt{n}}\right). \tag{31}$$

Taking the limit of $\epsilon' \uparrow \epsilon$, we see that for any arbitrary $[\tilde{R}, R - \tilde{R}]^\mathsf{T}$ in $\mathcal{Q}^{-1}(\mathbb{V}_{\mathsf{GP}}, \epsilon)$, we can achieve $P[\mathcal{E}] \leq \epsilon$. This completes the proof. ∎

C. Broadcast channel

Consider the problem of transmission of private messages over a broadcast channel depicted in Fig. 2. Let $q_{Y_1 Y_2 | X}$ be the channel transition probability.

1) One-shot achievability:

Definition 5. *An* $(\mathsf{M}_0, \mathsf{M}_1, \mathsf{M}_2)$-*code for broadcast channel consists of a (possibly stochastic) encoder* $\varphi : [1 : \mathsf{M}_0] \times [1 : \mathsf{M}_1] \times [1 : \mathsf{M}_2] \mapsto \mathcal{X}$, *and two (possibly stochastic) decoders* $\psi_k : \mathcal{Y}_k \mapsto [1 : \mathsf{M}_0] \times [1 : \mathsf{M}_k], k = 1, 2$.

Theorem 4. *Given any* q_{U_0, U_1, U_2} *and function* $x(u_0, u_1, u_2)$, *there is an* $(\mathsf{M}_0, \mathsf{M}_1, \mathsf{M}_2)$-*code for a single use of the channel whose probability of correct decoding is bounded from below by*

$$\mathbb{E}\left[\left(1 + (\mathsf{J}_1 \mathsf{J}_2)^{-1} 2^{\imath(U_1; U_2 | U_0)}\right) \prod_{k=1}^{2} \left(1 + \mathsf{M}_k \mathsf{J}_k 2^{-\imath(U_k; Y_k | U_0)} + \mathsf{M}_0 \mathsf{J}_k \mathsf{M}_k 2^{-\imath(U_0 U_k; Y_k)}\right)\right]^{-1},$$

where $J_1, J_2 > 0$ *are arbitrary integers. Moreover, loosening this bound gives the following upper bound on error probability of the code,*

$$\mathsf{P}\big[\log J_1 J_2 - \imath(U_1; U_2|U_0) < \gamma \ , \qquad \text{or}$$

$$\imath(U_1; Y_1|U_0) - \log \mathsf{M}_1 \mathsf{J}_1 < \gamma \quad , \text{or} \quad \imath(U_0 U_1; Y_1) - \log \mathsf{M}_0 \mathsf{M}_1 \mathsf{J}_1 < \gamma \text{ or}$$

$$\imath(U_2; Y_2|U_0) - \log \mathsf{M}_2 \mathsf{J}_2 < \gamma \quad , \text{or} \quad \imath(U_0 U_2; Y_2) - \log \mathsf{M}_0 \mathsf{M}_2 \mathsf{J}_2 < \gamma\big] + 17 \times 2^{-\gamma}, \tag{32}$$

where γ is any positive number.

Corollary 1. *Given any q_{U_1, U_2} and function $x(u_1, u_2)$, there is an $(\mathsf{M}_1, \mathsf{M}_2)$-code for a single use of the channel whose probability of correct decoding is bounded from below by*

$$\mathbb{E}\Bigg[\Big(1 + (\mathsf{J}_1 \mathsf{J}_2)^{-1} 2^{\imath(U_1; U_2)}\Big) \prod_{k=1}^{2} (1 + \mathsf{M}_k \mathsf{J}_k 2^{-\imath(U_k; Y_k)})\Bigg]^{-1},$$

where $J_1, J_2 > 0$ are arbitrary integers. Moreover, loosening this bound gives the following upper bound on error probability of the code,

$$\mathsf{P}\big[\log J_1 J_2 - \imath(U_1; U_2) < \gamma \ , \qquad \text{or}$$

$$\imath(U_1; Y_1) - \log \mathsf{M}_1 \mathsf{J}_1 < \gamma \ , \text{ or}$$

$$\imath(U_2; Y_2) - \log \mathsf{M}_2 \mathsf{J}_2 < \gamma\big] + 7 \times 2^{-\gamma}, \tag{33}$$

where γ is any positive number.

Remark 3. *Verdu derives a one-shot bound for the same problem in [9]. He derives the bound by proposing a one-shot covering and packing lemma. However to get access to the boundary of Marton's inner bound one needs a mutual covering lemma (since time sharing is not possible in one-shot and not useful in finite block length regime). For this reason Verdu's result seems to be weaker than ours. Our technique allows us to bypass the need for developing a one-shot version of the mutual covering lemma.*

Proof: For simplicity we prove the one-shot version of Marton with two auxiliaries (we will put the full proof in the next version of this draft). We only show the lower bound on probability of correct decoding. Derivation of the loosened bound is similar to that of Gelfand-Pinsker and thus omitted.

Random codebook generation: Let $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 = \{U_1(m_1, j_1)\}_{m_1=1, j_1=1}^{\mathsf{M}_1, \mathsf{J}_1} \times \{U_2(m_2, j_2)\}_{m_2=1, j_2=1}^{\mathsf{M}_2, \mathsf{J}_2}$ be a random product codebook, in which the codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ are generated independently and the elements of the codebook $\mathcal{C}_k, k = 1, 2$ are drawn independently from q_{U_k}. Thus the codebook is generated according to $r_{U_1 U_2} = q_{U_1} q_{U_2} \neq q_{U_1 U_2}$. Here J_1, J_2 are introducing redundancy but since it will be decoded at the receiver we can view these as dummy messages.

Encoding: Instead of using conventional mutual covering, we use an SMC which acts as follows. Given m_1, m_2, the SMC chooses indices j_1, j_2 with the probability

$$P_{\mathsf{Enc}}(j_1, j_2|m_1, m_2) = \frac{2^{\imath_q(U_1(m_1, j_1); U_2(m_2, j_2))}}{\sum_{\tilde{j}_1, \tilde{j}_2} 2^{\imath_q(U_1(m_1, \tilde{j}_1); U_2(m_2, \tilde{j}_2))}}. \tag{34}$$

Then the encoder transmits $x(U_1(m_1, j_1), U_2(m_2, j_2))$ through the channel. Observe that we generate codewords according to $r_{U_1 U_2}$ but compute the informations $\imath_q$ using $q_{U_1 U_2}$. This resembles the Marton coding scheme where we generate U_1^n and U_2^n independently but choose the jointly typical ones for transmission.

Decoding: We again use an SMC for decoding. Observing y_k, decoder k uses the following SMC to find both the message m_k and the dummy message j_k:

$$P_{\mathsf{Dec}_k}(\hat{m}_k, \hat{j}_k | y) = \frac{2^{\imath_q(y_k; U_k(\hat{m}_k, \hat{j}_k))}}{\sum_{\bar{m}_k, \bar{j}_k} 2^{\imath_q(y_k; U_k(\bar{m}_k, \bar{j}_k))}}.$$

Analysis: Observe that the joint distribution of random variables factors as,

$$P(m_{1:2}, j_{1:2}, y_{1:2}, \hat{m}_{1:2}, \hat{j}_{1:2}) = \frac{1}{\mathsf{M}_1 \mathsf{M}_2} P_{\mathsf{Enc}}(j_{1:2} | m_{1:2})$$

$$q(y_{1:2} | U_1(m_1, j_1), U_2(m_2, j_2)) \prod_{k=1}^{2} P_{\mathsf{Dec}_k}(\hat{m}_k, \hat{j}_k | y_k),$$

where $q(y_{1:2} | U_1(m_1, j_1), U_2(m_2, j_2))$ is equal to $q_{Y_{1:2}|X}(y_{1:2} | x(U_1(m_1, j_1), U_2(m_2, j_2)))$. We make an error if either of the decoders fail. The probability of correct decoding can be bounded from below by $\mathsf{P}[C] \geq \sum_{m_{1:2}, j_{1:2}, y_1} P(m_{1:2}, j_{1:2}, y_1, y_2, \hat{M}_{1:2} = m_{1:2}, \hat{J}_{1:2} = j_{1:2})$, hence skipping similar symmetry arguments we have:

$$\mathbb{E}\mathsf{P}[C] \geq \mathbb{E} \sum_{y_1, y_2} \mathsf{J}_1 \mathsf{J}_2 \frac{2^{\imath_q(U_1(1,1); U_2(1,1))}}{\sum_{\tilde{j}_1, \tilde{j}_2} 2^{\imath_q(U_1(1, \tilde{j}_1); U_2(1, \tilde{j}_2))}} q(y_1 | U_1(1,1), U_2(1,1)) \prod_{k=1}^{2} \frac{2^{\imath_q(y_k; U_k(1,1))}}{\sum_{\bar{m}_k, \bar{j}_k} 2^{\imath_q(y_k; U_k(\bar{m}_k, \bar{j}_k))}} \tag{35}$$

$$\geq \sum_{y_1, y_2} \mathbb{E}_{U_{1:2}(1,1)} \left(\frac{\mathsf{J}_1 \mathsf{J}_2 2^{\imath_q(U_1(1,1); U_2(1,1))}}{\mathbb{E}_{\mathcal{C}|U_{1:2}(1,1)} \sum_{\tilde{j}_1, \tilde{j}_2} 2^{\imath_q(U_1(1, \tilde{j}_1); U_2(1, \tilde{j}_2))}} \right.$$

$$\left. q(y_1 | U_{1:2}(1,1)) \prod_{k=1}^{2} \frac{2^{\imath_q(y_k; U_k(1,1))}}{\mathbb{E}_{\mathcal{C}|U_{1:2}(1,1)} \sum_{\bar{m}_k, \bar{j}_k} 2^{\imath_q(y_k; U_k(\bar{m}_k, \bar{j}_k))}} \right) \tag{36}$$

$$\geq \sum_{y_1, y_2} \mathbb{E}_{U_{1:2}(1,1)} \left(\frac{\mathsf{J}_1 \mathsf{J}_2 2^{\imath_q(U_1(1,1); U_2(1,1))}}{2^{\imath_q(U_1(1,1); U_2(1,1))} + \mathsf{J}_1 \mathsf{J}_2} q(y_1 | U_{1:2}(1,1)) \prod_{k=1}^{2} \frac{2^{\imath_q(y_k; U_k(1,1))}}{2^{\imath_q(y_k; U_k(1,1))} + \mathsf{M}_k \mathsf{J}_k} \right) \tag{37}$$

$$= \sum_{u_1, u_2, y_1, y_2} q(u_1, u_2, y_1) \frac{\mathsf{J}_1 \mathsf{J}_2}{2^{\imath_q(u_1; u_2)} + \mathsf{J}_1 \mathsf{J}_2} \prod_{k=1}^{2} \frac{2^{\imath_q(y_k; u_k)}}{2^{\imath_q(y_k; u_k)} + \mathsf{M}_k \mathsf{J}_k} \tag{38}$$

$$= \mathbb{E}_q \frac{1}{(1 + (\mathsf{J}_1 \mathsf{J}_2)^{-1} 2^{\imath_q(U_1; U_2)}) \prod_{k=1}^{2} (1 + \mathsf{M}_k \mathsf{J}_k 2^{-\imath_q(U_k; Y_k)})} \tag{39}$$

where (35) is due to symmetry, the <u>main step</u> (36) follows from Jensen inequality for the three-valued convex function $f(x_1, x_2, x_3) = \dfrac{1}{x_1 x_2 x_3}$ on $\mathbb{R}^3_+$. The expectation in (36) is over $U_{1:2}(1,1)$ of the codebook generation distributed according to $r_{U_1 U_2}$. Equation (37) follows from the following equations:

$$\mathbb{E}_{\mathcal{C}|U_{1:2}(1,1)} 2^{\imath_q(U_1(1,1); U_2(1, \tilde{j}_2))} = \sum_{u_2} q(u_2) 2^{\imath_q(U_1(1,1); u_2)}$$

$$= \sum_{u_2} q(u_2 | U(1,1)) = 1, \ \tilde{j}_2 \neq 1, \tag{40}$$

$$\mathbb{E}_{\mathcal{C}|U_{1:2}(1,1)} 2^{\imath_q(U_1(1, \tilde{j}_1); U_2(1,1))} = 1, \ \tilde{j}_1 \neq 1, \tag{41}$$

$$\mathbb{E}_{\mathcal{C}|U_{1:2}(1,1)} 2^{\imath_q(U_1(1,\tilde{j}_1);U_2(1,\tilde{j}_2))} = 1, \ \tilde{j}_1 \neq 1 \ \tilde{j}_2 \neq 1, \tag{42}$$

where in (40) we use the fact that $U_2(1,\tilde{j}_2)$ is independent of $(U_1(1,1), U_2(1,1))$ for $\tilde{j}_2 \neq 1$ and generated according to q_{U_2}. (41) and (42) follows similarly. Finally and (38) follows from the fact that $(U_1(1,1), U_2(1,1))$ is generated according to product distribution $r_{U_1U_2} = q_{U_1}q_{U_2}$. $\blacksquare$

Remark 4. *Unlike the Gelfand-Pinsker problem, the SMC encoder used is* (43) *cannot be written in the form of a SLC encoder. The SLC encoder has the following form:*

$$P_{\mathsf{Enc}}^{SLC}(j_1, j_2|m_1, m_2) = \frac{q(U_1(m_1, j_1), U_2(m_2, j_2))}{\sum_{\tilde{j}_1, \tilde{j}_2} q(U_1(m_1, \tilde{j}_1), U_2(m_2, \tilde{j}_2))}.$$

We use SMC instead of SLC, because the analysis of SLC is challenging.

2) Second Order achievability of broadcast channel:

Theorem 5. *Given a memoryless broadcast channel $q_{Y_1, Y_2|X}$, for any $(q_{U_1 U_2}, x(u_1, u_2))$, the pair (R_1, R_2) is (n, ϵ)-achievable, if there exists $(\tilde{R}_1, \tilde{R}_2)$ such that*

$$\begin{bmatrix} \tilde{R}_1 + \tilde{R}_2 \\ -R_1 - \tilde{R}_1 \\ -R_2 - \tilde{R}_2 \end{bmatrix} \in \mathbb{I}_{\mathsf{BC}} + \frac{1}{\sqrt{n}} \mathcal{Q}^{-1}(\mathbb{V}_{\mathsf{BC}}, \epsilon) + O\left(\frac{\log n}{n}\right) \tag{43}$$

where

$$\mathbb{I}_{\mathsf{BC}} = \mathbb{E}_{U_1 U_2 Y_1 Y_2} I_{\mathsf{BC}}, \qquad \mathbb{V}_{\mathsf{BC}} = \mathsf{Cov}(I_{\mathsf{BC}}), \tag{44}$$

in which

$$I_{\mathsf{BC}} = \begin{bmatrix} \imath(U_1; U_2) \\ -\imath(U_1; Y_1) \\ -\imath(U_2; Y_2) \end{bmatrix}. \tag{45}$$

Proof: The proof uses (33) in a way similar to the proof of Theorem 3 and hence is omitted. $\blacksquare$

IV. MULTI-TERMINAL LOSSY SOURCE CODING PROBLEMS

To illustrate the application of our technique to multi-terminal lossy source coding problems, we study the problems of Berger-Tung, Heegard-Berger/Kaspi and Multiple-description in this section. Throughout this section we use s for source and $\hat{s}$ for its reconstruction. Since $\hat{s}$ is created from a set of rv's available at the decoder, we follow El Gamal and Kim's notation [12] to also use $\hat{s}$ as a function of the rv's available at the decoder. For instance if the decoder has rv Y, we use a decoding function $\hat{s}(y)$.
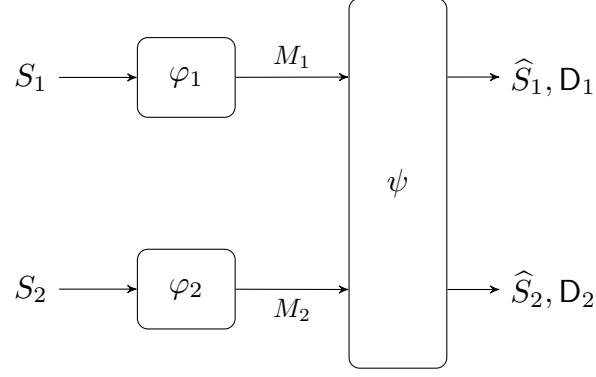
Fig. 3. Distributed lossy compression system

A. Berger-Tung

Consider the problem of distributed lossy compression depicted in Fig. 3. Let $q_{S_1 S_2}$ be the joint distribution of the sources and $\mathsf{d}_k(s_k, \hat{s}_k), k = 1, 2$, be two distortion measures. We will use the *probability of excess distortion* as the criterion for measuring the reliability of the system.

Definition 6. *An* $(\mathsf{M}_1, \mathsf{M}_2)$*-code for distributed lossy compression consists of (possibly stochastic) encoders* $\varphi_k :$ $\mathcal{S}_k \mapsto [1 : \mathsf{M}_k], k = 1, 2$*, and (possibly stochastic) decoder* $\psi : [1 : \mathsf{M}_1] \times [1 : \mathsf{M}_2] \mapsto \widehat{S}_1 \times \widehat{S}_2$*. Given a pair of distortion levels* $(\mathsf{D}_1, \mathsf{D}_2)$*, the probability of excess distortion is defined by,*

$$\mathsf{P}[\mathcal{E}; \mathsf{D}_1, \mathsf{D}_2] := \mathsf{P}[\mathsf{d}_1(S_1, \widehat{S}_1) > \mathsf{D}_1 \ \lor \ \mathsf{d}_2(S_2, \widehat{S}_2) > \mathsf{D}_2].$$

Also, we define the probability of no-excess distortion *by*

$$\mathsf{P}[C; \mathsf{D}_1, \mathsf{D}_2] := \mathsf{P}[\mathsf{d}_1(S_1, \widehat{S}_1) \leq \mathsf{D}_1, \ \mathsf{d}_2(S_2, \widehat{S}_2) \leq \mathsf{D}_2].$$

We prove a one-shot version of the result of Berger-Tung [13], [14] for this problem.

Theorem 6. *Given any pmf* $q_{U_1|S_1} q_{U_2|S_2}$ *and functions* $\hat{s}_k(u_1, u_2), k = 1, 2$*, there is an* $(\mathsf{M}_1, \mathsf{M}_2)$*-code for a single use of the sources whose probability of* no-excess distortion *is bounded from below by*

$$\mathsf{P}[C; \mathsf{D}_1, \mathsf{D}_2] \geq \mathbb{E} \frac{\mathbf{1}\left\{\mathsf{d}_1\left(S_1, \hat{s}_1(U_1, U_2)\right) \leq \mathsf{D}_1, \mathsf{d}_2\left(S_2, \hat{s}_2(U_1, U_2)\right) \leq \mathsf{D}_2\right\}}{(1 + \mathsf{J}_1^{-1} 2^{\imath(S_1; U_1)})(1 + \mathsf{J}_2^{-1} 2^{\imath(S_2; U_2)})\left(1 + (\mathsf{J}_2 \mathsf{M}_2^{-1} + \mathsf{J}_1 \mathsf{M}_1^{-1} + \mathsf{J}_1 \mathsf{J}_2 (\mathsf{M}_1 \mathsf{M}_2)^{-1}) 2^{-\imath(U_1; U_2)}\right)},$$
(46)

where $\mathsf{J}_k \geq \mathsf{M}_k, k = 1, 2$ *are arbitrary integers. Moreover, loosening this bound gives the following upper bound on the probability of excess distortion of the code,*

$$\mathsf{P}\left[\mathsf{d}_1\left(S_1, \hat{s}_1(U_1, U_2)\right) > \mathsf{D}_1, \text{or } \mathsf{d}_2\left(S_2, \hat{s}_2(U_1, U_2)\right) > \mathsf{D}_2, \text{or}\right.$$

$$\left. \log \mathsf{J}_1 - \imath(S_1; U_1) < \gamma, \text{or} \ \log \mathsf{J}_2 - \imath(S_2; U_2) < \gamma, \text{or} \ \imath(U_1; U_2) - \log \frac{\mathsf{J}_1 \mathsf{J}_2}{\mathsf{M}_1 \mathsf{M}_2} < \gamma\right] + 15 \times 2^{-\gamma}. \quad (47)$$

Remark 5. *The term* $1 + \mathsf{J}_k^{-1} 2^{\imath(S_k; U_k)}, \quad k = 1, 2$ *corresponds to a covering of* S_k *with* U_k *in the asymptotic regime. The alphabet size of* U_k *is* J_k*. We use a random binning of* U_k*, mapping it from a set of size* J_k *to a set of size* M_k*. This explains the inequality* $\mathsf{J}_k \geq \mathsf{M}_k$ *for* $k = 1, 2$*. The term* $\left(1 + (\mathsf{J}_2 \mathsf{M}_2^{-1} + \mathsf{J}_1 \mathsf{M}_1^{-1} + \mathsf{J}_1 \mathsf{J}_2 (\mathsf{M}_1 \mathsf{M}_2)^{-1}) 2^{-\imath(U_1; U_2)}\right)$ *corresponds to a mutual packing lemma in the asymptotic regime.*

Proof: We only prove the lower bound on probability of no-excess distortion. Derivation of the loosened bound is similar to that of Gelfand-Pinsker and is thus omitted.

Random codebook generation: Let $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2 = \{U_1(j_1)\}_{j_1=1}^{\mathsf{J}_1} \times \{U_2(j_2)\}_{j_2=1}^{\mathsf{J}_2}$ be a random product codebook, in which the codebooks $\mathcal{C}_1$ and $\mathcal{C}_2$ are generated independently and the elements of the codebook $\mathcal{C}_k, k = 1, 2$ are drawn independently from q_{U_k}. Thus the codebook is generated according to $r_{U_1 U_2} = q_{U_1} q_{U_2} \neq q_{U_1 U_2}$.

Random binning: Let $\mathcal{B}_k : [1 : \mathsf{J}_k] \mapsto [1 : \mathsf{M}_k], k = 1, 2$ be two independent random mappings (binning), in which $\mathcal{B}_k$ maps each element of $[1 : \mathsf{J}_k]$ uniformly and independently to the set $[1 : \mathsf{M}_k]$.

Encoding: Encoder $k = 1, 2$ uses an SMC followed by a random binning to obtain the index m_k. First given s_k, the SMC chooses an index j_k with the probability

$$P_{\mathsf{Enc}_k}(j_k | s_k) = \frac{2^{\imath_q(s_k; U(j_k))}}{\sum_{\tilde{j}_k} 2^{\imath_q(s_k; U(\tilde{j}_k))}}.$$

Then the encoder transmits $m_k = \mathcal{B}_k(j_k)$ to the decoder.

Decoding: We use a modified SMC for decoding w.r.t. a receiving indices (m_1, m_2). Observing (m_1, m_2), decoder uses the following modified SMC to find both the j_1 and j_2 and thus (U_1, U_2):

$$P_{\mathsf{Dec}}(\hat{j}_1, \hat{j}_2 | m_1, m_2) = \frac{2^{\imath_q(U_1(\hat{j}_1); U_2(\hat{j}_2))} \mathbf{1}\{B_1(\hat{j}_1) = m_1, B_2(\hat{j}_2) = m_2\}}{\sum_{\bar{j}_1, \bar{j}_2} 2^{\imath_q(U_1(\bar{j}_1); U_2(\bar{j}_2))} \mathbf{1}\{B_1(\bar{j}_1) = m_1, B_2(\bar{j}_2) = m_2\}}.$$

Remark: Observe that the above SMC resembles a joint-typical decoder for the mutual packing lemma in the asymptotic regime. It can be considered as a dual of the SMC encoder of equation (43) that corresponded to a mutual covering lemma in the asymptotic regime.

Given m_1 and m_2, the decoder chooses among the pairs $(U_1(j_1), U_2(j_2))$ assigned to the bin (m_1, m_2). The higher the information between $U_1(j_1)$ and $U_2(j_2)$, the more likely we choose it at the decoder. Finally, decoder computes $\hat{s}_k(U_1(\hat{j}_1), U_2(\hat{j}_2))$ as the estimate of S_k.

Analysis: Observe that the joint distribution of random variables factors as,

$$P(s_{1:2}, m_{1:2}, j_{1:2}, \hat{j}_{1:2}) = q(s_1, s_2) \prod_{k=1}^{2} P_{\mathsf{Enc}_k}(j_k | s_k) \mathbf{1}\{B_1(j_1) = m_1, B_2(j_2) = m_2\} P_{\mathsf{Dec}}(\hat{j}_1, \hat{j}_2 | m_1, m_2).$$

We consider two error events:

1) The decoder fails to recover the correct pair (j_1, j_2), i.e. $(\hat{j}_1, \hat{j}_2) \neq (j_1, j_2)$.
2) One of the distortions corresponding to the pair (j_1, j_2) exceeds the designated levels, i.e. $\mathsf{d}_k\big(s_k, \hat{s}_k(U_1(j_1), U_2(j_2))\big) > \mathsf{D}_k$ for some $k \in \{1, 2\}$.

The probability of correct decoding can be bounded from below by

$$\mathsf{P}[C] \geq \sum_{s_{1:2}, m_{1:2}, j_{1:2}} P(s_{1:2}, m_{1:2}, j_{1:2}, \hat{J}_1 = j_1, \hat{J}_2 = j_2)$$

$$\times \mathbf{1}\left\{\mathsf{d}_1\left(s_1,\hat{s}_1(U_1(j_1),U_2(j_2))\right)\leq \mathsf{D}_1,\mathsf{d}_2\left(s_2,\hat{s}_2(U_1(j_1),U_2(j_2))\right)\leq \mathsf{D}_2\right\},$$

hence skipping similar symmetry arguments we have:

$$\mathbb{E}\mathsf{P}[C]\geq \mathbb{E}_{\mathcal{C},\mathcal{B}}\sum_{s_1,s_2}\mathsf{M}_1\mathsf{M}_2\mathsf{J}_1\mathsf{J}_2 q(s_1,s_2)\frac{2^{\imath(s_1;U_1(1))}}{\sum_{\tilde{j}_1}2^{\imath(s_1;U_1(\tilde{j}_1))}}\mathbf{1}(B_1(1)=1)\frac{2^{\imath(s_2;U_2(1))}}{\sum_{\tilde{j}_2}2^{\imath(s_2;U_2(\tilde{j}_2))}}\mathbf{1}(B_2(1)=1)$$

$$\times \frac{2^{\imath(U_1(1);U_2(1))}\mathbf{1}(B_1(1)=1,B_2(1)=1)}{\sum_{\bar{j}_1,\bar{j}_2}2^{\imath(U_1(\bar{j}_1);U_2(\bar{j}_2))}\mathbf{1}\{B_1(\bar{j}_1)=1,B_2(\bar{j}_2)=1\}}$$

$$\times \mathbf{1}\left\{\mathsf{d}_1\left(s_1,\hat{s}_1(U_1(1),U_2(1))\right)\leq \mathsf{D}_1,\mathsf{d}_2\left(s_2,\hat{s}_2(U_1(1),U_2(1))\right)\leq \mathsf{D}_2\right\} \tag{48}$$

$$\geq \mathsf{M}_1\mathsf{M}_2\sum_{s_1,s_2}q(s_1,s_2)\mathbb{E}_{U_{1:2}(1),B_{1:2}(1)}\left(\frac{\mathsf{J}_1 2^{\imath(s_1;U_1(1))}}{\mathbb{E}_{\mathcal{C}|U_{1:2}(1)}\sum_{\tilde{j}_1}2^{\imath(s_1;U_1(\tilde{j}_1))}}\frac{\mathsf{J}_2 2^{\imath(s_2;U_2(1))}}{\mathbb{E}_{\mathcal{C}|U_{1:2}(1)}\sum_{\tilde{j}_2}2^{\imath(s_2;U_2(\tilde{j}_2))}}\right.$$

$$\times \frac{2^{\imath(U_1(1);U_2(1))}\mathbf{1}(B_1(1)=1,B_2(1)=1)}{\mathbb{E}_{\mathcal{C},\mathcal{B}|U_{1:2}(1),B_{1:2}(1)}\sum_{\bar{j}_1,\bar{j}_2}2^{\imath(U_1(\bar{j}_1);U_2(\bar{j}_2))}\mathbf{1}\{B_1(\bar{j}_1)=1,B_2(\bar{j}_2)=1\}}$$

$$\times \left. \mathbf{1}\left\{\mathsf{d}_1\left(s_1,\hat{s}_1(U_1(1),U_2(1))\right)\leq \mathsf{D}_1,\mathsf{d}_2\left(s_2,\hat{s}_2(U_1(1),U_2(1))\right)\leq \mathsf{D}_2\right\}\right) \tag{49}$$

$$\geq \mathsf{M}_1\mathsf{M}_2\sum_{s_1,s_2}q(s_1,s_2)\mathbb{E}_{U_{1:2}(1),B_{1:2}(1)}\left(\frac{2^{\imath(s_1;U_1(1))}}{1+\mathsf{J}_1^{-1}2^{\imath(s_1;U_1(1))}}\frac{2^{\imath(s_2;U_2(1))}}{1+\mathsf{J}_2^{-1}2^{\imath(s_2;U_2(1))}}\right.$$

$$\times \frac{2^{\imath(U_1(1);U_2(1))}\mathbf{1}(B_1(1)=1,B_2(1)=1)}{2^{\imath(U_1(1);U_2(1))}\mathbf{1}(B_1(1)=1,B_2(1)=1)+\mathsf{J}_2\mathsf{M}_2^{-1}\mathbf{1}(B_1(1)=1)+\mathsf{J}_1\mathsf{M}_1^{-1}\mathbf{1}(B_2(1)=1)+\mathsf{J}_1\mathsf{J}_2(\mathsf{M}_1\mathsf{M}_2)^{-1}}$$

$$\times \left. \mathbf{1}\left\{\mathsf{d}_1\left(s_1,\hat{s}_1(U_1(1),U_2(1))\right)\leq \mathsf{D}_1,\mathsf{d}_2\left(s_2,\hat{s}_2(U_1(1),U_2(1))\right)\leq \mathsf{D}_2\right\}\right) \tag{50}$$

$$= \sum_{s_1,s_2,u_1,u_2}q(s_1,s_2)q(u_1|s_1)q(u_2|s_2)\frac{1}{1+\mathsf{J}_1^{-1}2^{\imath(s_1;u_1)}}\frac{1}{1+\mathsf{J}_2^{-1}2^{\imath(s_2;u_2)}}$$

$$\times \frac{\mathbf{1}\left\{\mathsf{d}_1\left(s_1,\hat{s}_1(u_1,u_2)\right)\leq \mathsf{D}_1,\mathsf{d}_2\left(s_2,\hat{s}_2(u_1,u_2)\right)\leq \mathsf{D}_2\right\}}{1+(\mathsf{J}_2\mathsf{M}_2^{-1}+\mathsf{J}_1\mathsf{M}_1^{-1}+\mathsf{J}_1\mathsf{J}_2(\mathsf{M}_1\mathsf{M}_2)^{-1})2^{-\imath(u_1;u_2)}} \tag{51}$$

$$= \mathbb{E}_{q_{S_{1:2}U_{1:2}}}\frac{\mathbf{1}\left\{\mathsf{d}_1\left(S_1,\hat{s}_1(U_1,U_2)\right)\leq \mathsf{D}_1,\mathsf{d}_2\left(S_2,\hat{s}_2(U_1,U_2)\right)\leq \mathsf{D}_2\right\}}{(1+\mathsf{J}_1^{-1}2^{\imath(S_1;U_1)})(1+\mathsf{J}_2^{-1}2^{\imath(S_2;U_2)})\left(1+(\mathsf{J}_2\mathsf{M}_2^{-1}+\mathsf{J}_1\mathsf{M}_1^{-1}+\mathsf{J}_1\mathsf{J}_2(\mathsf{M}_1\mathsf{M}_2)^{-1})2^{-\imath(U_1;U_2)}\right)} \tag{52}$$

where (48) is due to symmetry, the <u>main step</u> (49) follows from Jensen inequality for the three-valued convex function $f(x_1,x_2,x_3)=\dfrac{1}{x_1 x_2 x_3}$ on $\mathbb{R}_+^3$. The expectation in (49) is over $U_{1:2}(1,1)$ of the codebook generation distributed according to $r_{U_1 U_2}$. Equation (50) follows from separating the cases $(\bar{j}_1=1,\bar{j}_2\neq 1)$, $(\bar{j}_1\neq 1,\bar{j}_2=1)$, $(\bar{j}_1\neq 1,\bar{j}_2\neq 1)$ and $(\bar{j}_1=1,\bar{j}_2=1)$. This is discussed below. Lastly, equation (52) follows from the facts that random codebook and random binning are independent and $(U_1(1),U_2(1))$ is generated according to product distribution $r_{U_1 U_2}=q_{U_1}q_{U_2}$.

Derivation of equation (50): In the following we consider each of the four case mentioned above separately.

- Case 1, $(\bar{j}_1=1,\bar{j}_2\neq 1)$:

$$\mathbb{E}_{\mathcal{C},\mathcal{B}|U_{1:2}(1),B_{1:2}(1)}2^{\imath(U_1(1);U_2(\bar{j}_2))}\mathbf{1}\{B_1(1)=1,B_2(\bar{j}_2)=1\}=\sum_{u_2}q(u_2)2^{\imath(U_1(1);u_2)}\mathbf{1}\{B_1(1)=1\}\mathsf{M}_2^{-1}$$

$$= \mathsf{M}_2^{-1}\sum_{u_2}q(u_2|U_1(1))\mathbf{1}\{B_1(1)=1\}=\mathsf{M}_2^{-1}\mathbf{1}\{B_1(1)=1\} \tag{53}$$
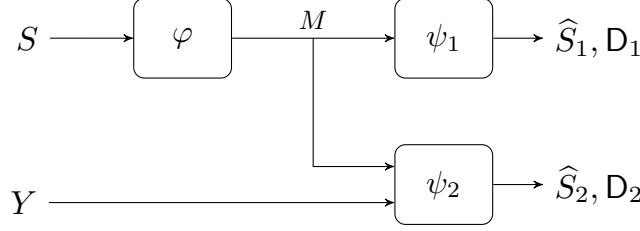
where in (53) we use the following facts:

Fig. 4. Lossy source coding when side information is absent.

1) $U_2(\bar{j}_2)$ is independent of $(U_1(1), U_2(1))$ for $\bar{j}_2 \neq 1$ and generated according to q_{U_2}.

2) The random binning and random codebook are independent and $B_2(\bar{j}_2)$ has a uniform distribution over $[1 : M_2]$.

- Case 2, $(\bar{j}_1 \neq 1, \bar{j}_2 = 1)$: This is similar to the previous case and gives us

$$\mathbb{E}_{\mathcal{C}, \mathcal{B}|U_{1:2}(1), B_{1:2}(1)} 2^{\imath(U_1(\bar{j}_1); U_2(1))} \mathbf{1}\{B_1(\bar{j}_1) = 1, B_2(1) = 1\} = M_1^{-1} \mathbf{1}\{B_2(1) = 1\} \tag{54}$$

- Case 3: $(\bar{j}_1 \neq 1, \bar{j}_2 \neq 1)$: In this case $U_1(\bar{j}_1), U_2(\bar{j}_2)$ are independent of $U_{1:2}(1), B_{1:2}(1)$, and are generated according to $q_{U_1} q_{U_2}$.

$$\mathbb{E}_{\mathcal{C}, \mathcal{B}|U_{1:2}(1), B_{1:2}(1)} 2^{\imath(U_1(\bar{j}_1); U_2(\bar{j}_2))} \mathbf{1}\{B_1(\bar{j}_1) = 1, B_2(\bar{j}_2) = 1\} = \sum_{u_1, u_2} q(u_1) q(u_2) 2^{\imath(u_1; u_2)} (M_1 M_2)^{-1}$$

$$= (M_1 M_2)^{-1} \sum_{u_1, u_2} q(u_1, u_2) = (M_1 M_2)^{-1}. \tag{55}$$

- Case 4: $(\bar{j}_1 = 1, \bar{j}_2 = 1)$:

$$\mathbb{E}_{\mathcal{C}, \mathcal{B}|U_{1:2}(1), B_{1:2}(1)} 2^{\imath(U_1(1); U_2(1))} \mathbf{1}\{B_1(1) = 1, B_2(1) = 1\} = 2^{\imath(U_1(1); U_2(1))} \mathbf{1}\{B_1(1) = 1, B_2(1) = 1\}. \tag{56}$$

■

B. Heegard-Berger/Kaspi

Consider the problem of lossy source coding when side information may be absent, depicted in Fig. 4. Let q_S be the distribution of the source and $d_k(s, \hat{s}_k), k = 1, 2$, be two distortion measures. We will use the *probability of excess distortion* as the criterion for measuring the reliability of the system.

Definition 7. *An M-code for distributed lossy compression consists of (possibly stochastic) encoder $\varphi : \mathcal{S} \mapsto [1 : M]$, and (possibly stochastic) decoders $\psi_1 : [1 : M] \mapsto \widehat{\mathcal{S}}_1$ and $\psi_2 : [1 : M] \times \mathcal{Y} \mapsto \widehat{\mathcal{S}}_2$. Given a pair of distortion levels (D_1, D_2), the probability of excess distortion is defined by,*

$$P[\mathcal{E}; D_1, D_2] := P[d_1(S, \widehat{S}_1) > D_1 \ \vee \ d_2(S, \widehat{S}_2) > D_2].$$

Also, we define the probability of no-excess distortion *by*

$$P[C; D_1, D_2] := P[d_1(S, \widehat{S}_1) \leq D_1, \ d_2(S, \widehat{S}_2) \leq D_2].$$

Remark 6. *Setting* $D_1 = \infty$, *this problem reduces to the Wyner-Ziv problem. A one-shot result for the Wyner-Ziv problem has been concurrently obtained by [10].*

We prove a one-shot version of the result of Heegard-Berger/Kaspi [15], [16] for this problem.

Theorem 7. *Given any* $q_{WU|S}$ *and functions* $\hat{s}_1(w), \hat{s}_2(w, u, y)$, *there is an* M-*code for a single use of the source whose probability of* no-excess distortion *is bounded from below by*

$$P[C; D_1, D_2] \geq \mathbb{E} \frac{\mathbf{1}\{d_1(S, \hat{s}_1(W)) \leq D_1, d_2(S, \hat{s}_2(W, U, Y)) \leq D_2\}}{\left(1 + M_1^{-1} 2^{\imath(S;W)} + (M_1 J_2)^{-1} 2^{\imath(S;W,U)}\right)\left(1 + J_2 M_2^{-1} 2^{-\imath(Y;U|W)}\right)}, \tag{57}$$

where M_1, M_2, J_2 *are integers such that* $M = M_1 M_2$ *and* $J_2 \geq M_2$. *Moreover, loosening this bound gives the following upper bound on the probability of excess distortion of the code,*

$$P\left[d_1(S, \hat{s}_1(W)) > D_1, \text{or } d_2(S, \hat{s}_2(W, U, Y)) > D_2, \text{or}\right.$$

$$\left.\log M_1 - \imath(S;W) < \gamma, \text{or} \quad \log M_1 J_2 - \imath(S;WU) < \gamma, \text{or} \quad \imath(Y;U|W) - \log J_2 M_2^{-1} < \gamma\right] + 5 \times 2^{-\gamma}. \tag{58}$$

Proof: We only prove the lower bound on probability of no-excess distortion. Derivation of the loosened bound is similar to that of Gelfand-Pinsker and is thus omitted.

Random codebook generation:

- $\mathcal{C}_1 = \{W(m_1)\}_{m_1=1}^{M_1}$ is a random codebook whose elements are drawn independently from q_W.
- For each $m_1 \in [1 : M_1]$, let $\mathcal{C}_2(m_1) = \{U(m_1, j_2)\}_{j_2=1}^{J_2}$ be a random codebook whose elements are drawn independently from $q_{U|W}(.|W(m_1))$. Moreover the codebooks $\mathcal{C}_2(m_1), m_1 \in [1 : M_1]$ are generated independently.
- Let $\mathcal{C}$ be the set of all codewords.

Random binning: Let $\mathcal{B} : [1 : J_2] \mapsto [1 : M_2]$ be a random mapping (binning), where $\mathcal{B}$ maps each element of $[1 : J_2]$ uniformly and independently to the set $[1 : M_2]$.

Encoding: Encoder uses an SMC followed by a random binning to obtain the pair (m_1, m_2). First given s, the SMC chooses a pair (m_1, j_2) with the probability

$$P_{\mathsf{Enc}}(m_1, j_2|s) = \frac{2^{\imath_q(s;W(m_1), U(m_1, j_2))}}{\sum_{\tilde{m}_1, \tilde{j}_2} 2^{\imath_q(s;W(\tilde{m}_1), U(\tilde{m}_1, \tilde{j}_2))}}.$$

Then the encoder computes $m_2 = \mathcal{B}(j_2)$ and sends $m = (m_1, m_2)$ to the decoders.

Decoding: Decoder 1 outputs $\hat{s}_1(W(m_1))$ as the estimate of S. Decoder 2 uses a modified SMC for decoding w.r.t. a received index (m_1, m_2). Observing (m_1, m_2), decoder 2 uses the following modified SMC to find j_2 and thus $(W(m_1), U(m_1, j_2))$:

$$P_{\mathsf{Dec}}(\hat{j}_2|m_1, m_2) = \frac{2^{\imath_q(y;U(m_1, \hat{j}_2)|W(m_1))} \mathbf{1}(B(\hat{j}_2) = m_2)}{\sum_{\bar{j}_2} 2^{\imath_q(y;U(m_1, \bar{j}_2)|W(m_1))} \mathbf{1}(B(\bar{j}_2) = m_2)}.$$

Observe that the above SMC resembles a joint-typical decoder of the asymptotic regime. Finally, decoder 2 computes $\hat{s}_2(W(m_1), U(m_1, \hat{j}_2), y)$ as the estimate of S.

Analysis: Observe that the joint distribution of random variables factors as,

$$P(s, y, m_{1:2}, j_2, \hat{j}_2) = q(s,y)P_{\mathsf{Enc}}(m_1, j_2|s)\mathbf{1}\{B(\hat{j}_2) = m_2\}P_{\mathsf{Dec}}(\hat{j}_2|m_1, m_2).$$

We consider two error events:

1) Decoder 2 fails to recover the correct j_2, i.e. $\hat{j}_2 \neq j_2$.

2) The distortion of decoder 1, or that of decoder 2 corresponding to j_2 exceeds the designated levels, i.e.
   $\mathsf{d}_1\big(s, \hat{s}_1(W(m_1))\big) > \mathsf{D}_1$ or $\mathsf{d}_2\big(s, \hat{s}_2(W(m_1), U(m_1, \hat{j}_2), y)\big) > \mathsf{D}_2$.

The probability of correct decoding can be bounded from below by

$$\mathsf{P}[C] \geq \sum_{s,y,m_{1:2},j_2} P(s, y, m_{1:2}, j_2, \hat{J}_2 = j_2)$$

$$\times \mathbf{1}\left\{\mathsf{d}_1\left(s, \hat{s}_1(W(m_1))\right) \leq \mathsf{D}_1, \mathsf{d}_2\left(s, \hat{s}_2(W(m_1), U(m_1, j_2), y)\right) \leq \mathsf{D}_2\right\},$$

hence skipping similar symmetry arguments we have:

$$\mathbb{E}\mathsf{P}[C] \geq \mathbb{E}_{\mathcal{C},\mathcal{B}} \sum_{s,y} \mathsf{M}_1\mathsf{M}_2\mathsf{J}_2 q(s,y) \frac{2^{i(s;W(1),U(1,1))}}{\sum_{\tilde{m}_1,\tilde{j}_2} 2^{i(s;W(\tilde{m}_1),U(\tilde{m}_1,\tilde{j}_2))}}\mathbf{1}(B(1) = 1)\frac{2^{i(y;U(1,1)|W(1))}\mathbf{1}(B(1) = 1)}{\sum_{\bar{j}_2} 2^{i(y;U(1,\bar{j}_2)|W(1))}\mathbf{1}(B(\bar{j}_2) = 1)}$$

$$\times \mathbf{1}\left\{\mathsf{d}_1\left(s, \hat{s}_1(W(1))\right) \leq \mathsf{D}_1, \mathsf{d}_2\left(s, \hat{s}_2(W(1), U(1,1), y)\right) \leq \mathsf{D}_2\right\} \tag{59}$$

$$\geq \mathsf{M}_2 \sum_{s,y} q(s,y)\mathbb{E}_{W(1),U(1,1),B(1)}\left(\frac{\mathsf{M}_1\mathsf{J}_2 2^{i(s;W(1),U(1,1))}}{\mathbb{E}_{\mathcal{C}|W(1),U(1,1)} \sum_{\tilde{m}_1,\tilde{j}_2} 2^{i(s;W(\tilde{m}_1),U(\tilde{m}_1,\tilde{j}_2))}}\right.$$

$$\times \frac{2^{i(y;U(1,1)|W(1))}\mathbf{1}(B(1) = 1)}{\mathbb{E}_{\mathcal{C},\mathcal{B}|W(1),U(1,1),B(1)} \sum_{\bar{j}_2} 2^{i(y;U(1,\bar{j}_2)|W(1))}\mathbf{1}(B(\bar{j}_2) = 1)}$$

$$\left.\times \mathbf{1}\left\{\mathsf{d}_1\left(s, \hat{s}_1(W(1))\right) \leq \mathsf{D}_1, \mathsf{d}_2\left(s, \hat{s}_2(W(1), U(1,1), y)\right) \leq \mathsf{D}_2\right\}\right) \tag{60}$$

$$\geq \mathsf{M}_2 \sum_{s,y} q(s,y)\mathbb{E}_{W(1),U(1,1),B(1)}\left(\frac{\mathsf{M}_1\mathsf{J}_2 2^{i(s;W(1),U(1,1))}}{\mathsf{M}_1\mathsf{J}_2 + \mathsf{J}_2 2^{i(s;W(1))} + 2^{i(s;W(1),U(1,1))}}\right.$$

$$\times \frac{2^{i(y;U(1,1)|W(1))}\mathbf{1}(B(1) = 1)}{2^{i(y;U(1,1)|W(1))}\mathbf{1}(B(1) = 1) + \mathsf{J}_2\mathsf{M}_2^{-1}}$$

$$\left.\times \mathbf{1}\left\{\mathsf{d}_1\left(s, \hat{s}_1(W(1))\right) \leq \mathsf{D}_1, \mathsf{d}_2\left(s, \hat{s}_2(W(1), U(1,1), y)\right) \leq \mathsf{D}_2\right\}\right) \tag{61}$$

$$= \sum_{s,y,w,u} q(s,y)q(w,u|s)\left(\frac{1}{1 + \mathsf{M}_1^{-1}2^{i(s;w)} + (\mathsf{M}_1\mathsf{J}_2)^{-1}2^{i(s;w,u)}} \cdot \frac{1}{1 + \mathsf{J}_2\mathsf{M}_2^{-1}2^{-i(y;u|w)}}\right.$$

$$\left.\times \mathbf{1}\left\{\mathsf{d}_1\left(s, \hat{s}_1(w)\right) \leq \mathsf{D}_1, \mathsf{d}_2\left(s, \hat{s}_2(w, u, y)\right) \leq \mathsf{D}_2\right\}\right) \tag{62}$$

$$= \mathbb{E}\frac{\mathbf{1}\left\{\mathsf{d}_1\left(S, \hat{s}_1(W)\right) \leq \mathsf{D}_1, \mathsf{d}_2\left(S, \hat{s}_2(W, U, Y)\right) \leq \mathsf{D}_2\right\}}{\left(1 + \mathsf{M}_1^{-1}2^{i(S;W)} + (\mathsf{M}_1\mathsf{J}_2)^{-1}2^{i(S;W,U)}\right)\left(1 + \mathsf{J}_2\mathsf{M}_2^{-1}2^{-i(Y;U|W)}\right)} \tag{63}$$

where (59) is due to symmetry, the <u>main step</u> (60) follows from Jensen inequality for the two-valued convex function $f(x_1, x_2) = \dfrac{1}{x_1 x_2}$ on $\mathbb{R}_+^2$. Equation (61) follows from computing the denominator terms. This is discussed below. Lastly, equation (62) follows from the facts that random codebook and random binning are independent and $(W(1), U(1,1))$ is generated according to q_{WU}.
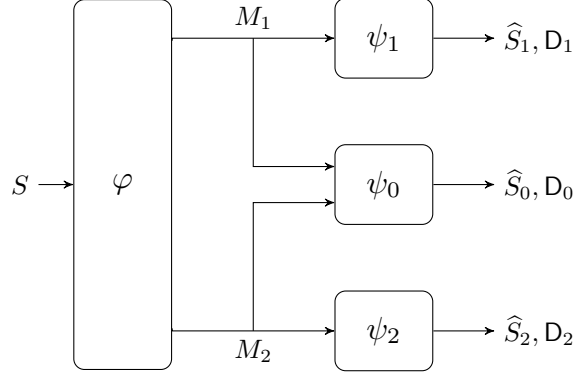
Fig. 5. Multiple description coding.

Derivation of equation (61)*:* Deriving this equation involves computing two denominator terms. To compute $\mathbb{E}_{\mathcal{C},\mathcal{B}|W(1),U(1,1),B(1)} \sum_{\bar{j}_2} 2^{\imath(y;U(1,\bar{j}_2)|W(1))} \mathbf{1}(B(\bar{j}_2) = 1)$ for any $\bar{j}_2 \neq 1$ we have:

$$\mathbb{E}_{\mathcal{C},\mathcal{B}|W(1),U(1,1),B(1)} 2^{\imath(y;U(1,\bar{j}_2)|W(1))} \mathbf{1}(B(\bar{j}_2) = 1) = \sum_u q(u|W(1)) 2^{\imath(y;U(1,\bar{j}_2)|W(1))} \mathsf{M}_2^{-1}$$

$$= \mathsf{M}_2^{-1} \sum_u q(u|W(1),y) = \mathsf{M}_2^{-1}, \tag{64}$$

where we have used similar argument used in the proof of (53).

To compute $\mathbb{E}_{\mathcal{C}|W(1),U(1,1)} \sum_{\tilde{m}_1,\tilde{j}_2} 2^{\imath(s;W(\tilde{m}_1),U(\tilde{m}_1,\tilde{j}_2))}$ we consider the following three cases separately:

- Case 1, $(\tilde{m}_1 = 1, \tilde{j}_2 \neq 1)$:

$$\mathbb{E}_{\mathcal{C}|W(1),U(1,1)} 2^{\imath(s;W(1),U(1,\tilde{j}_2))} = \sum_u q(u|W(1)) 2^{\imath(s;W(1),u)} = \sum_u \frac{q(W(1),u|s)}{q(W(1))} = 2^{\imath(s;W(1))}, \tag{65}$$

  where we have used the fact that $U(1,\tilde{j}_2), \tilde{j}_2 \neq 1$ is independent of $W(1), U(1,1)$ and drawn from $q_{U|W}(.|W(1))$.

- Case 2, $\tilde{m}_1 \neq 1$:

$$\mathbb{E}_{\mathcal{C}|W(1),U(1,1)} 2^{\imath(s;W(\tilde{m}_1),U(\tilde{m}_1,\tilde{j}_2))} = 1, \tag{66}$$

  where we have used the fact that $U(\tilde{m}_1,\tilde{j}_2), \tilde{m}_1 \neq 1$ is independent of $W(1), U(1,1)$ and drawn from q_{WU}.

- Case 3, $(\tilde{m}_1 = 1, \tilde{j}_2 = 1)$:

$$\mathbb{E}_{\mathcal{C}|W(1),U(1,1)} 2^{\imath(s;W(1),U(1,1)} = 2^{\imath(s;W(1),U(1,1)}. \tag{67}$$

∎

C. Multiple Description

Consider the problem of multiple description coding depicted in Fig. 5. Let q_S be the distribution of the source and $\mathsf{d}_k(s, \hat{s}_k), k = 0, 1, 2$ be three distortion measures. We will use the *probability of excess distortion* as the criterion for measuring the reliability of the system.

Definition 8. *An* (M_1, M_2)-*code for the multiple description coding consists of (possibly stochastic) encoder* $\varphi :$ $\mathcal{S} \mapsto [1 : M_1] \times [1 : M_2]$, *and (possibly stochastic) decoders* $\psi_0 : [1 : M_1] \times [1 : M_2] \mapsto \widehat{S}_0$, $\psi_1 : [1 : M_1] \mapsto \widehat{S}_1$ *and* $\psi_2 : [1 : M_2] \mapsto \widehat{S}_2$. *Given a tuple of distortion levels* (D_0, D_1, D_2), *the probability of excess distortion is defined by,*

$$P[\mathcal{E}; D_0, D_1, D_2] := P[d_0(S, \widehat{S}_0) > D_0 \ \vee \ d_1(S, \widehat{S}) > D_1 \ \vee \ d_2(S, \widehat{S}_2) > D_2].$$

Also, we define the probability of no-excess distortion *by*

$$P[C; D_0, D_1, D_2] := P[d_0(S, \widehat{S}_0) \leq D_0, d_1(S, \widehat{S}_1) \leq D_1, \ d_2(S, \widehat{S}_2) \leq D_2].$$

We prove a one-shot version of an equivalent characterization of Zhang-Berger inner bound [17], [18] for this problem.

Theorem 8. *Given any* $q_{U_0 U_1 U_2|S}$ *and functions* $\hat{s}_0(u_0, u_1, u_2)$, $\hat{s}_1(u_0, u_1)$ *and* $\hat{s}_2(u_0, u_2)$, *there is an* (M_1, M_2)-*code for a single use of the source whose probability of* no-excess distortion *is bounded from below by*

$$P[C; D_0, D_1, D_2] \geq \mathbb{E} \frac{\mathbf{1} \left\{ d_0 \left(S, \hat{s}_0(U_0, U_1, U_2)\right) \leq D_0, d_1 \left(S, \hat{s}_1(U_0, U_1)\right) \leq D_1, d_2 \left(S, \hat{s}_2(U_0, U_2)\right) \leq D_2 \right\}}{1 + J_0^{-1} 2^{\imath(S; U_0)} + M_1^{-1} 2^{\imath(S; U_0 U_1)} + M_2^{-1} 2^{\imath(S; U_0 U_2)} + J_0(M_1 M_2)^{-1} 2^{\imath(S; U_0 U_1 U_2) + \imath(U_1; U_2|U_0)}},$$
(68)

where J_0 *is a common divisor of* M_1 *and* M_2. *Moreover, loosening this bound gives the following upper bound on the probability of excess distortion of the code,*

$$P \left[d_0 \left(S, \hat{s}_0(U_0, U_1, U_2)\right) > D_0, \text{or } d_1 \left(S, \hat{s}_1(U_0, U_1)\right) > D_1, \text{or } d_2 \left(S, \hat{s}_2(U_0, U_2)\right) > D_2, \text{or} \right.$$

$$\log J_0 - \imath(S; U_0) < \gamma, \text{or } \log M_1 - \imath(S; U_0 U_1) < \gamma, \text{or } \log M_2 - \imath(S; U_0 U_2) < \gamma, \text{or}$$

$$\left. \log \frac{M_1 M_2}{J_0} - \imath(S; U_0 U_1 U_2) - \imath(U_1; U_2|U_0) < \gamma \right] + 4 \times 2^{-\gamma}.$$
(69)

Proof: We only prove the lower bound on probability of no-excess distortion. Derivation of the loosened bound is similar to that of Gelfand-Pinsker and thus omitted. Since J_0 is a common divisor of M_1 and M_2, there exists positive integers J_1 and J_2 such that $M_1 = J_0 J_1$ and $M_2 = J_0 J_2$.

Random codebook generation:

- $\mathcal{C}_0 = \{U_0(j_0)\}_{j_0=1}^{J_0}$ is a random codebook whose elements are drawn independently from q_{U_0}.
- For each $j_0 \in [1 : J_0]$, let $\mathcal{C}_{12}(j_0) = \mathcal{C}_1(j_0) \times \mathcal{C}_2(j_0) = \{U_1(j_1)\}_{j_1=1}^{J_1} \times \{U_2(j_2)\}_{j_2=1}^{J_2}$ be a random product codebook, in which the codebooks $\mathcal{C}_1(j_0)$ and $\mathcal{C}_2(j_0)$ are generated independently and the elements of the codebook $\mathcal{C}_k(j_0), k = 1, 2$ are drawn independently from $q_{U_k|U_0}(.|U_0(j_0))$. Moreover the codebooks $\mathcal{C}_{12}(j_0), j_0 \in [1 : J_0]$ are generated independently.
- Let $\mathcal{C}$ be the set of all codewords.

Encoding: Encoder uses an SMC for three r.v.'s to obtain the tuple (j_0, j_1, j_2). Given s, the SMC chooses a tuple (j_0, j_1, j_2) with the probability

$$P_{\mathsf{Enc}}(j_0, j_1, j_2|s) = \frac{2^{\imath_q(s; U_0(j_0), U_1(j_0, j_1), U_2(j_0, j_2)) + \imath_q(U_1(j_0, j_1); U_2(j_0, j_2)|U_0(1))}}{\sum_{\tilde{j}_0, \tilde{j}_1, \tilde{j}_2} 2^{\imath_q(s; U_0(\tilde{j}_0), U_1(\tilde{j}_0, \tilde{j}_1), U_2(\tilde{j}_0, \tilde{j}_2)) + \imath_q(U_1(\tilde{j}_0, \tilde{j}_1); U_2(\tilde{j}_0, \tilde{j}_2)|U_0(\tilde{j}_0))}}.$$

Then the encoder sends $m_1 = (j_0, j_1)$ and $m_2 = (j_0, j_2)$ to the decoders.

Decoding: Observing $(m_1, m_2) = (j_0, j_1, j_2)$, decoder 0 computes $\hat{s}_0(U_0(j_0), U_1(j_0, j_1), U_2(j_0, j_2))$ as the estimate of S. Observing $m_k = (j_0, j_k)$, decoder $k = 1, 2$ computes $\hat{s}_0(U_0(j_0), U_k(j_0, j_k))$ as the estimate of S.

Analysis: Observe that the joint distribution of random variables factors as,

$$P(s, j_0, j_1, j_2) = q(s) P_{\mathsf{Enc}}(j_0, j_1, j_2 | s).$$

The probability of no-excess distortion can be bounded from below by

$$\mathsf{P}[C] \geq \sum_{s, y, m_{1:2}, j_2} P(s, j_0, j_1, j_2) \chi_{s, U_0(j_0), U_1(j_0, j_1), U_2(j_0, j_2)},$$

where

$$\chi_{s, u_0, u_1, u_2} := \mathbf{1}\left\{\mathsf{d}_0\left(S, \hat{s}_0(u_0, u_1, u_2)\right) > \mathsf{D}_0, \mathsf{d}_1\left(S, \hat{s}_1(u_0, u_1)\right) > \mathsf{D}_1, \mathsf{d}_2\left(S, \hat{s}_2(u_0, u_2)\right) > \mathsf{D}_2\right\}.$$

Skipping similar symmetry arguments we have:

$$\mathbb{E}\, \mathsf{P}[C] \geq \mathbb{E}_{\mathcal{C}} \sum_s \mathsf{J}_0 \mathsf{J}_1 \mathsf{J}_2 q(s) \frac{2^{\imath(s; U_0(1), U_1(1,1), U_2(1,1)) + \imath(U_1(1,1); U_2(1,1) | U_0(1))}}{\sum_{\tilde{j}_0, \tilde{j}_1, \tilde{j}_2} 2^{\imath(s; U_0(\tilde{j}_0), U_1(\tilde{j}_0, \tilde{j}_1), U_2(\tilde{j}_0, \tilde{j}_2)) + \imath(U_1(\tilde{j}_0, \tilde{j}_1); U_2(\tilde{j}_0, \tilde{j}_2) | U_0(\tilde{j}_0))}} \chi_{s, U_0(1), U_1(1,1), U_2(1,1)}$$

$$\geq \sum_s q(s) \mathbb{E}_{U_{0:2}(1,1,1)} \frac{\mathsf{J}_0 \mathsf{J}_1 \mathsf{J}_2 2^{\imath(s; U_0(1), U_1(1,1), U_2(1,1)) + \imath(U_1(1,1); U_2(1,1) | U_0(1))} \chi_{s, U_0(1), U_1(1,1), U_2(1,1)}}{\mathbb{E}_{\mathcal{C} | U_0(1), U_1(1,1), U_2(1,1)} \sum_{\tilde{j}_0, \tilde{j}_1, \tilde{j}_2} 2^{\imath(s; U_0(\tilde{j}_0), U_1(\tilde{j}_0, \tilde{j}_1), U_2(\tilde{j}_0, \tilde{j}_2)) + \imath(U_1(\tilde{j}_0, \tilde{j}_1); U_2(\tilde{j}_0, \tilde{j}_2) | U_0(\tilde{j}_0))}} \tag{70}$$

$$\geq \sum_s q(s) \mathbb{E}_{U_{0:2}(1,1,1)} \mathsf{J}_0 \mathsf{J}_1 \mathsf{J}_2 2^{\imath(s; U_0(1), U_1(1,1), U_2(1,1)) + \imath(U_1(1,1); U_2(1,1) | U_0(1))} \chi_{s, U_0(1), U_1(1,1), U_2(1,1)}$$

$$\times \left(\mathsf{J}_0 \mathsf{J}_1 \mathsf{J}_2 + \mathsf{J}_1 \mathsf{J}_2 2^{\imath(s; U_0(1))} + \mathsf{J}_2 2^{\imath(s; U_0(1), U_1(1,1))} + \mathsf{J}_1 2^{\imath(s; U_0(1), U_2(1,1))}\right.$$

$$\left. + 2^{\imath(s; U_0(1), U_1(1,1), U_2(1,1)) + \imath(U_1(1,1); U_2(1,1) | U_0(1))}\right)^{-1} \tag{71}$$

$$= \sum_{s, u_0, u_1, u_2} q(s, u_0, u_1, u_2) \frac{\chi_{s, u_0, u_1, u_2}}{1 + \mathsf{J}_0^{-1} 2^{\imath(s; u_0)} + (\mathsf{J}_0 \mathsf{J}_1)^{-1} 2^{\imath(s; u_0 u_1)} + (\mathsf{J}_0 \mathsf{J}_2)^{-1} 2^{\imath(s; u_0 u_2)} + (\mathsf{J}_0 \mathsf{J}_1 \mathsf{J}_2)^{-1} 2^{\imath(s; u_0 u_1 u_2) + \imath(u_1; u_2 | u_0)}} \tag{72}$$

$$= \mathbb{E} \frac{\mathbf{1}\left\{\mathsf{d}_0\left(S, \hat{s}_0(U_0, U_1, U_2)\right) > \mathsf{D}_0, \mathsf{d}_1\left(S, \hat{s}_1(U_0, U_1)\right) > \mathsf{D}_1, \mathsf{d}_2\left(S, \hat{s}_2(U_0, U_2)\right) > \mathsf{D}_2\right\}}{1 + \mathsf{J}_0^{-1} 2^{\imath(S; U_0)} + \mathsf{M}_1^{-1} 2^{\imath(S; U_0 U_1)} + \mathsf{M}_2^{-1} 2^{\imath(S; U_0 U_2)} + \mathsf{J}_0(\mathsf{M}_1 \mathsf{M}_2)^{-1} 2^{\imath(S; U_0 U_1 U_2) + \imath(U_1; U_2 | U_0)}} \tag{73}$$

where the <u>main step</u> (70) follows from Jensen inequality for the two-valued convex function $f(x) = \dfrac{1}{x}$ on $\mathbb{R}_+$. Equation (71) follows from separating the cases ..... This is discussed below. Lastly, equation (72) follows from the facts that $(U_0(1), U_1(1,1), U_2(1,1))$ is generated according to $q_{U_0} q_{U_1 | U_0} q_{U_2 | U_0}$.

Derivation of equation (71)*:* In the following we consider each of the three case mentioned above separately. But first observe that

$$2^{\imath(s; u_0, u_1, u_2) + \imath(u_1; u_2 | u_0)} = \frac{q(u_0, u_1, u_2)}{q(u_0) q(u_1 | u_0) q(u_2 | u_0)}.$$

- Case 1, $\tilde{j}_0 \neq 1$:

$$\mathbb{E}_{\mathcal{C} | U_0(1), U_1(1,1), U_2(1,1)} 2^{\imath(s; U_0(\tilde{j}_0), U_1(\tilde{j}_0, \tilde{j}_1), U_2(\tilde{j}_0, \tilde{j}_2)) + \imath(U_1(\tilde{j}_0, \tilde{j}_1); U_2(\tilde{j}_0, \tilde{j}_2) | U_0(\tilde{j}_0))}$$

$$= \sum_{u_0, u_1, u_2} q(u_0) q(u_1 | u_0) q(u_2 | u_0) 2^{\imath(s; u_0, u_1, u_2) + \imath(u_1; u_2 | u_0)}$$

$$= 1, \tag{74}$$

where we have used the fact that $(U_0(\tilde{j}_0), U_1(\tilde{j}_0, \tilde{j}_1), U_2(\tilde{j}_0, \tilde{j}_2)), \tilde{j}_0 \neq 1$ is independent of $(U_0(1), U_1(1,1), U_2(1,1))$ and drawn from $q_{U_0} q_{U_1|U_0} q_{U_2|U_0}$.

- Case 2, $\tilde{j}_0 = 1$ and $\tilde{j}_1 \neq 1, \tilde{j}_2 \neq 1$:

$$\mathbb{E}_{\mathcal{C}|U_0(1), U_1(1,1), U_2(1,1)} 2^{\imath(s; U_0(1), U_1(1, \tilde{j}_1), U_2(1, \tilde{j}_2)) + \imath(U_1(1, \tilde{j}_1); U_2(1, \tilde{j}_2)|U_0(1))}$$

$$= \sum_{u_1, u_2} q(u_1|U_0(1)) q(u_2|U_0(1)) 2^{\imath(s; U_0(1), u_1, u_2) + \imath(u_1; u_2|U_0(1))}$$

$$= \sum_{u_1, u_2} \frac{q(U_0(1), u_1, u_2|s)}{q(U_0(1))}$$

$$= 2^{\imath(s; U_0(1))}, \tag{75}$$

where we have used the fact that given $U_0(1)$, $(U_1(1, \tilde{j}_1), U_2(1, \tilde{j}_2)), \tilde{j}_1 \neq 1, \tilde{j}_2 \neq 1$ is independent of $(U_1(1,1), U_2(1,1))$ and drawn from $q_{U_1|U_0}(.|U_0(1)) q_{U_2|U_0}(.|U_0(1))$.

- Case 3, $\tilde{j}_0 = \tilde{j}_1 = 1$ and $\tilde{j}_2 \neq 1$: Similarly we have,

$$\mathbb{E}_{\mathcal{C}|U_0(1), U_1(1,1), U_2(1,1)} 2^{\imath(s; U_0(1), U_1(1,1), U_2(1, \tilde{j}_2)) + \imath(U_1(1,1); U_2(1, \tilde{j}_2)|U_0(1))}$$

$$= 2^{\imath(s; U_0(1), U_1(1,1))} \tag{76}$$

- Case 4, $\tilde{j}_0 = \tilde{j}_2 = 1$ and $\tilde{j}_1 \neq 1$: Similarly we have,

$$\mathbb{E}_{\mathcal{C}|U_0(1), U_1(1,1), U_2(1,1)} 2^{\imath(s; U_0(1), U_1(1, \tilde{j}_1), U_2(1,1)) + \imath(U_1(1, \tilde{j}_1); U_2(1,1)|U_0(1))}$$

$$= 2^{\imath(s; U_0(1), U_2(1,1))} \tag{77}$$

- Case 5, $(\tilde{j}_0, \tilde{j}_1, \tilde{j}_2) = (1, 1, 1)$:

$$\mathbb{E}_{\mathcal{C}|U_0(1), U_1(1,1), U_2(1,1)} 2^{\imath(s; U_0(1), U_1(1,1), U_2(1,1)) + \imath(U_1(1,1); U_2(1,1)|U_0(1))}$$

$$= 2^{\imath(s; U_0(1), U_1(1,1), U_2(1,1)) + \imath(U_1(1,1); U_2(1,1)|U_0(1))}. \tag{78}$$

Applying (74)-(77) to (70) yields (71). ∎

V. Joint source-channel coding

To illustrate the application of our technique to joint source-channel coding problems, we study the problem of joint source-channel coding over MAC in this section.

A. Joint source-channel coding over MAC

Consider the problem of transmission of correlated sources (S_1, S_2) over a MAC depicted in Fig. 6. Let $q_{Y|X_1 X_2}$ be channel transition probability. We prove a one-shot version of the result of Cover-El Gamal-Salehi [19] for this problem.
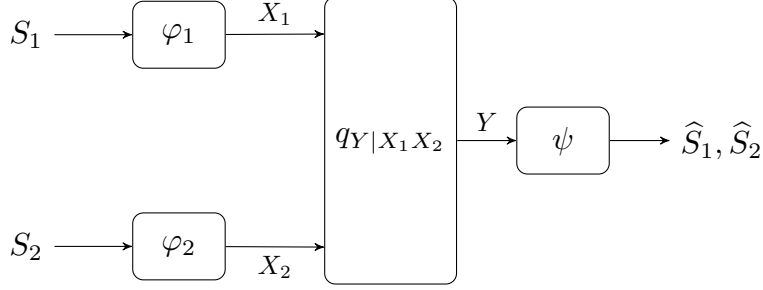
Fig. 6. Joint source-channel coding over MAC.

Definition 9. *An code for lossless transmission of correlated source over a MAC consists of (possibly stochastic) encoders $\varphi_k : \mathcal{S}_k \mapsto \mathcal{X}_k, k = 1, 2$, and (possibly stochastic) decoder $\psi : \mathcal{Y} \mapsto \mathcal{S}_1 \times \mathcal{S}_2$.*

Theorem 9. *Let K be the common part of S_1 and S_2. Given any $q(s_1, s_2, t, x_1, x_2) = q(s_1, s_2)q(t)q(x_1|s_1, t)q(x_2|s_2, t)$, there is a code for a single use of the channel whose probability of correct decoding is bounded from below by*

$$\mathbb{E}\left(1 + 2^{h(S_1|S_2) - \imath(Y; X_1|X_2, S_2, T)} + 2^{h(S_2|S_1) - \imath(Y; X_2|X_1, S_1, T)} + 2^{h(S_1, S_2|K) - \imath(Y; X_1, X_2|K, T)} + 2^{h(S_1, S_2) - \imath(Y; X_1, X_2)}\right)^{-1}.$$

Moreover, loosening this bound gives the following upper bound on error probability of the code,

$$\mathsf{P}[\imath(Y; X_1|X_2, S_2, T) - h(S_1|S_2) < \gamma , \quad \text{or}$$

$$\imath(Y; X_2|X_1, S_1, T) - h(S_2|S_1) < \gamma , \quad \text{or}$$

$$\imath(Y; X_1, X_2|K, T) - h(S_1, S_2|K) < \gamma , \quad \text{or}$$

$$\imath(Y; X_1, X_2) - h(S_1, S_2) < \gamma] + 4 \times 2^{-\gamma}, \tag{79}$$

where γ is any positive number.

Remark 7. *Second order analysis of the above bound can be obtained in a similar manner as the Gelfand-Pinsker rate region. We will include it in the next version of this draft. Its equations resemble the four constraints given by [19]. It subsumes the result of [6] whose region includes two additional constraints.*

Proof: We only prove the lower bound on probability of correct decoding. Derivation of the loosened bound is similar to that of Gelfand-Pinsker and is thus omitted.

Codebook generation: We employ a one-shot version of the codebook used in [19].

- For each $k \in \mathcal{K}$, draw a symbol $T(k)$ from pmf $q(t)$,
- For each s_1, draw a symbol $X_1(s_1)$ from $q(x_1|s_1, T(k(s_1)))$,
- For each s_2, draw a symbol $X_2(s_2)$ from $q(x_2|s_2, T(k(s_2)))$.

Encoding: Given s_j, transmitter j sends $X_j(s_j)$.

Decoding: In contrast to Gelfand-Pinsker and broadcast channel, we use an SLC for decoding. Observing y, decoder uses the SLC $P_{S_1 S_2 | Y}(\hat{s}_1, \hat{s}_2 | y)$ to find an estimate of (s_1, s_2). The SLC can be formulated as follows:

$$P_{S_1 S_2 | Y}(\hat{s}_1, \hat{s}_2 | y) = \frac{q(\hat{s}_1, \hat{s}_2) q(y | X_1(\hat{s}_1), X_2(\hat{s}_2))}{\sum_{\bar{s}_1, \bar{s}_2} q(\bar{s}_1, \bar{s}_2) q(y | X_1(\bar{s}_1), X_2(\bar{s}_2))}$$

$$= \frac{q(\hat{s}_1, \hat{s}_2) 2^{\imath_q(y; X_1(\hat{s}_1), X_2(\hat{s}_2))}}{\sum_{\bar{s}_1, \bar{s}_2} q(\bar{s}_1, \bar{s}_2) 2^{\imath_q(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}}.$$

Analysis: Observe that the joint distribution of random variables factors as,

$$P(s_1, s_2, y, \hat{s}_1, \hat{s}_2) = q(s_1, s_2) q(y | X_1(s_1), X_2(s_2)) P_{S_1 S_2 | Y}(\hat{s}_1, \hat{s}_2 | y).$$

The probability of correct decoding can be bounded from below by $\mathsf{P}[C] \geq \sum_{s_1, s_2, y} P(s_1, s_2, y, \hat{s}_1 = s_1, \hat{s}_2 = s_2)$, hence we have:

$$\mathbb{E}\mathsf{P}[C] = \mathbb{E} \sum_{s_1, s_2, y} q(s_1, s_2) q(y | (X_1(s_1), X_2(s_2)) \frac{q(s_1, s_2) 2^{\imath(y; X_1(s_1), X_2(s_2))}}{\sum_{\bar{s}_1, \bar{s}_2} q(\bar{s}_1, \bar{s}_2) 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}} \tag{80}$$

$$= \mathbb{E} \sum_{\substack{s_1, s_2, y: \\ (s_1, s_2) \in \mathsf{Supp}}} q(s_1, s_2) q(y | (X_1(s_1), X_2(s_2)) \frac{q(s_1, s_2) 2^{\imath(y; X_1(s_1), X_2(s_2))}}{\sum_{\bar{s}_1, \bar{s}_2} q(\bar{s}_1, \bar{s}_2) 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}} \tag{81}$$

$$\geq \sum_{\substack{s_1, s_2, y: \\ (s_1, s_2) \in \mathsf{Supp}}} q(s_1, s_2) \mathbb{E}_{T(k), X_1(s_1), X_2(s_2)} q(y | (X_1(s_1), X_2(s_2)) \frac{q(s_1, s_2) 2^{\imath(y; X_1(s_1), X_2(s_2))}}{\mathbb{E}_{\mathcal{C} | T(k), X_1(s_1), X_2(s_2)} \sum_{\bar{s}_1, \bar{s}_2} q(\bar{s}_1, \bar{s}_2) 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}} \tag{82}$$

where Supp is the support set of the pmf $q_{S_1 S_2}$, that is, $\mathsf{Supp} = \{(s_1, s_2) : q_{S_1 S_2}(s_1, s_2) \neq 0\}$. By $T(k)$ we mean $T(k(s_1)) = T(k(s_2))$, and the main step (82) follows from the Jensen inequality for the convex function $f(x) = \dfrac{1}{x}$ on $\mathbb{R}_+$. To evaluate (82), we first find an upper bound on the denominator. To do this, we first split the denominator into five terms as follows,

$$\mathbb{E}_{\mathcal{C} | T(k), X_1(s_1), X_2(s_2)} \sum_{\bar{s}_1, \bar{s}_2} q(\bar{s}_1, \bar{s}_2) 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))} = \sum_{\substack{(\bar{s}_1, \bar{s}_2) \in \mathsf{Supp}: \\ k(\bar{s}_1) = k(\bar{s}_2) \neq k}} q(\bar{s}_1, \bar{s}_2) \mathbb{E}_{\mathcal{C} | T(k), X_1(s_1), X_2(s_2)} 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}$$

$$+ \sum_{\substack{(\bar{s}_1, \bar{s}_2) \in \mathsf{Supp}: \\ \bar{s}_1 \neq s_1, \bar{s}_2 \neq s_2, k(\bar{s}_1) = k(\bar{s}_2) = k}} q(\bar{s}_1, \bar{s}_2) \mathbb{E}_{\mathcal{C} | T(k), X_1(s_1), X_2(s_2)} 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}$$

$$+ \sum_{\substack{\bar{s}_1 : (\bar{s}_1, s_2) \in \mathsf{Supp}: \\ \bar{s}_1 \neq s_1}} q(\bar{s}_1, s_2) \mathbb{E}_{\mathcal{C} | T(k), X_1(s_1), X_2(s_2)} 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}$$

$$+ \sum_{\substack{\bar{s}_2 : (s_1, \bar{s}_2) \in \mathsf{Supp}: \\ \bar{s}_2 \neq s_2}} q(s_1, \bar{s}_2) \mathbb{E}_{\mathcal{C} | T(k), X_1(s_1), X_2(s_2)} 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}$$

$$+ q(s_1, s_2) 2^{\imath(y; X_1(s_1), X_2(s_2))}. \tag{83}$$

We find upper bounds on each term of (83) separately.

- Case 1: $\bar{k} \neq k(= k(s_1) = k(s_2))$, where $\bar{k} = k(\bar{s}_1) = k(\bar{s}_2)$,

$$\mathbb{E}_{\mathcal{C} | T(k), X_1(s_1), X_2(s_2)} 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))} = \sum_{t, x_1, x_2} q(t) q(x_1 | \bar{s}_1, t) q(x_2 | \bar{s}_2, t) 2^{\imath(y; x_1, x_2)} \tag{84}$$

$$= \sum_{t, x_1, x_2} q(t, x_1, x_2 | \bar{s}_1, \bar{s}_2) \frac{q(y | x_1, x_2)}{q(y)} \tag{85}$$

$$= \frac{\sum_{t,x_1,x_2} q(t,x_1,x_2,y|\bar{s}_1,\bar{s}_2)}{q(y)} \tag{86}$$

$$= \frac{q(y|\bar{s}_1,\bar{s}_2)}{q(y)}, \tag{87}$$

where (84) and (85) follows from the fact that $(T(\bar{k}), X_1(\bar{s}_1), X_2(\bar{s}_2))$ is independent of $(T(k), X_1(s_1), X_2(s_2))$ and is drawn from $q_T q_{X_1|S_1 T}(.|\bar{s}_1,.) q_{X_2|S_2 T}(.|\bar{s}_2,.) = q_{T X_1 X_2|S_1,S_2}(.|\bar{s}_1,\bar{s}_2)$. Using (87) we obtain the following upper bound on the first term of (83):

$$\sum_{\substack{(\bar{s}_1,\bar{s}_2)\in\mathsf{Supp:}\\ \bar{k}\neq k}} q(\bar{s}_1,\bar{s}_2)\mathbb{E}_{\mathcal{C}|T(k),X_1(s_1),X_2(s_2)} 2^{\imath(y;X_1(\bar{s}_1),X_2(\bar{s}_2))} \leq \sum_{(\bar{s}_1,\bar{s}_2)\in\mathsf{Supp}} \frac{q(\bar{s}_1,\bar{s}_2,y)}{q(y)}$$

$$= 1. \tag{88}$$

- Case 2: $\bar{s}_1 \neq s_1$, $\bar{s}_2 \neq s_2$, but $\bar{k} = k(s_1) = k(s_2) = k$,

$$\mathbb{E}_{\mathcal{C}|T(k),X_1(s_1),X_2(s_2)} 2^{\imath(y;X_1(\bar{s}_1),X_2(\bar{s}_2))} = \sum_{x_1,x_2} q(x_1|\bar{s}_1,T(k))q(x_2|\bar{s}_2,T(k)) 2^{\imath(y;x_1,x_2)} \tag{89}$$

$$= \sum_{x_1,x_2} q(x_1,x_2|\bar{s}_1,\bar{s}_2,T(k))\frac{q(y|x_1,x_2)}{q(y)} \tag{90}$$

$$= \frac{\sum_{x_1,x_2} q(x_1,x_2,y|\bar{s}_1,\bar{s}_2,T(k))}{q(y)} \tag{91}$$

$$= \frac{q(y|\bar{s}_1,\bar{s}_2,T(k))}{q(y)}, \tag{92}$$

where (89) and (90) follows from the fact that given $T(k)$, $(X_1(\bar{s}_1), X_2(\bar{s}_2))$ is independent of $(X_1(s_1), X_2(s_2))$ and is drawn from $q_{X_1|S_1 T}(.|\bar{s}_1,T(k))q_{X_2|S_2 T}(.|\bar{s}_2,T(k)) = q_{X_1 X_2|S_1,S_2,T}(.|\bar{s}_1,\bar{s}_2,T(k))$. Using (92) we obtain the following upper bound on the second term of (83):

$$\sum_{\substack{(\bar{s}_1,\bar{s}_2)\in\mathsf{Supp:}\\ \bar{s}_1\neq s_1,\bar{s}_2\neq s_2,k(\bar{s}_1)=k}} q(\bar{s}_1,\bar{s}_2)\mathbb{E}_{\mathcal{C}|T(k),X_1(s_1),X_2(s_2)} 2^{\imath(y;X_1(\bar{s}_1),X_2(\bar{s}_2))} \leq \sum_{\substack{(\bar{s}_1,\bar{s}_2)\in\mathsf{Supp:}\\ k(\bar{s}_1)=k}} q(\bar{s}_1,\bar{s}_2|T(k))\frac{q(y|\bar{s}_1,\bar{s}_2,T(k))}{q(y)} \tag{93}$$

$$= \sum_{\substack{(\bar{s}_1,\bar{s}_2)\in\mathsf{Supp:}\\ k(\bar{s}_1)=k}} \frac{q(\bar{s}_1,\bar{s}_2,y|T(k))}{q(y)}$$

$$= \frac{q(k,y|T(k))}{q(y)} \tag{94}$$

$$= q(k)\frac{q(y|k,T(k))}{q(y)} = q(k)2^{\imath(y;k,T(k))} \tag{95}$$

where (93) is due to the fact that T is independent of (S_1, S_2) and (94) follows from the definition of common part, and $q(k) = \sum_{\substack{(\bar{s}_1,\bar{s}_2)\in\mathsf{Supp:}\\ k(\bar{s}_1)=k}} q(\bar{s}_1,\bar{s}_2)$.

- Case 3: $\bar{s}_1 \neq s_1$, $\bar{s}_2 = s_2$, hence $\bar{k} = k(s_2) = k$,

$$\mathbb{E}_{\mathcal{C}|T(k),X_1(s_1),X_2(s_2)} 2^{\imath(y;X_1(\bar{s}_1),X_2(s_2))} = \sum_{x_1} q(x_1|\bar{s}_1,T(k)) 2^{\imath(y;x_1,X_2(s_2))} \tag{96}$$

$$= \sum_{x_1} q(x_1|\bar{s}_1,s_2,T(k),X_2(s_2))\frac{q(y|x_1,X_2(s_2))}{q(y)} \tag{97}$$

$$= \frac{\sum_{x_1} q(x_1,y|\bar{s}_1,s_2,T(k),X_2(s_2))}{q(y)} \tag{98}$$

$$= \frac{q(y|\bar{s}_1, s_2, T(k), X_2(s_2))}{q(y)}, \qquad (99)$$

where (96) and (97) follows from the fact that given $(T(k), X_2(s_2))$, $X_1(\bar{s}_1)$ is independent of $X_1(s_1)$ and is drawn from $q_{X_1|S_1 T}(.|\bar{s}_1, T(k)) = q_{X_1 X_2|S_1, S_2, T}(.|\bar{s}_1, s_2, T(k), X_2(s_2))$. Using (99) we obtain the following upper bound on the third term of (83):

$$\sum_{\substack{\bar{s}_1:(\bar{s}_1, s_2)\in\mathsf{Supp} \\ \bar{s}_1 \neq s_1}} q(\bar{s}_1, s_2) \quad \mathbb{E}_{\mathcal{C}|T(k), X_1(s_1), X_2(s_2)} 2^{\imath(y; X_1(\bar{s}_1), X_2(s_2))}$$

$$\leq \sum_{\bar{s}_1:(\bar{s}_1, s_2)\in\mathsf{Supp}} q(s_2) q(\bar{s}_1|s_2, T(k)) \frac{q(y|\bar{s}_1, s_2, T(k), X_2(s_2))}{q(y)} \qquad (100)$$

$$\leq \sum_{\bar{s}_1:(\bar{s}_1, s_2)\in\mathsf{Supp}} q(s_2) q(\bar{s}_1|s_2, T(k), X_2(s_2)) \frac{q(y|\bar{s}_1, s_2, T(k), X_2(s_2))}{q(y)} \qquad (101)$$

$$\leq \sum_{\bar{s}_1} q(s_2) \frac{q(\bar{s}_1, y|s_2, T(k), X_2(s_2))}{q(y)} \qquad (102)$$

$$= q(s_2) \frac{q(y|s_2, T(k), X_2(s_2))}{q(y)} = q(s_2) 2^{\imath(y; s_2, T(k), X_2(s_2))}, \qquad (103)$$

where (100) is due to the fact that T is independent of (S_1, S_2) in the pmf q that we started with (this should be confused with the pmf induced by the code), and (101) follows from the Markov chain $S_1 - S_2 T - X_2$.

- Case 4: $\bar{s}_2 \neq s_2$, $\bar{s}_1 = s_1$, hence $k = k(s_1) = k$, Using symmetry between Case 3 and Case 4, we have the following bound on the fourth term of (83):

$$\sum_{\substack{\bar{s}_1:(s_1, \bar{s}_2)\in\mathsf{Supp} \\ \bar{s}_2 \neq s_2}} q(s_1, \bar{s}_2) \mathbb{E}_{\mathcal{C}|T(k), X_1(s_1), X_2(s_2)} 2^{\imath(y; X_1(s_1), X_2(\bar{s}_2))} \leq q(s_1) 2^{\imath(y; s_1, T(k), X_1(s_1))}. \qquad (104)$$

In summary, we have the following upper bound on the denominator of (82):

$$\mathbb{E}_{\mathcal{C}|T(k), X_1(s_1), X_2(s_2)} \sum_{\bar{s}_1, \bar{s}_2} q(\bar{s}_1, \bar{s}_2) 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))} \leq 1 + q(k) 2^{\imath(y; k, T(k))} + q(s_2) 2^{\imath(y; s_2, T(k), X_2(s_2))}$$

$$+ q(s_1) 2^{\imath(y; s_1, T(k), X_1(s_1))} + q(s_1, s_2) 2^{\imath(y; X_1(s_1), X_2(s_2))}. \qquad (105)$$

Substituting this in (82) gives:

$$\frac{q(s_1, s_2) 2^{\imath(y; X_1(s_1), X_2(s_2))}}{\mathbb{E}_{\mathcal{C}|T(k), X_1(s_1), X_2(s_2)} \sum_{\bar{s}_1, \bar{s}_2} q(\bar{s}_1, \bar{s}_2) 2^{\imath(y; X_1(\bar{s}_1), X_2(\bar{s}_2))}} \geq \Big(q(s_1, s_2)^{-1} 2^{-\imath(y; X_1(s_1), X_2(s_2))}$$

$$+ q(s_1, s_2|k)^{-1} 2^{\imath(y; k, T(k)) - \imath(y; X_1(s_1), X_2(s_2))}$$

$$+ q(s_1|s_2)^{-1} 2^{\imath(y; s_2, T(k), X_2(s_2)) - \imath(y; X_1(s_1), X_2(s_2))}$$

$$+ q(s_2|s_1)^{-1} 2^{\imath(y; s_1, T(k), X_1(s_1)) - \imath(y; X_1(s_1), X_2(s_2))} + 1 \Big)^{-1}$$

$$= \Big(2^{h(s_1, s_2) - \imath(y; X_1(s_1), X_2(s_2))}$$

$$+ 2^{h(s_1, s_2|k) - \imath(y; X_1(s_1), X_2(s_2)|k, T(k))}$$

$$+ 2^{h(s_1|s_2) - \imath(y; X_1(s_1)|s_2, T(k), X_2(s_2))}$$

$$+2^{h(s_2|s_1)-\imath(y;X_2(s_2)|s_1,T(k),X_1(s_1))}+1\Big)^{-1}. \qquad (106)$$

Using this and the fact that $(T(k), X_1(s_1), X_2(s_2))$ is drawn from $q_T q_{X_1|S_1 T}(.|s_1, .) q_{X_2|S_2 T}(.|s_2, .) = q_{T X_1 X_2|S_1,S_2}(.|s_1, s_2)$, we have:

$$\mathbb{E}\mathsf{P}[C] \geq \sum_{s_1,s_2,y} \sum_{t,x_1,x_2} q(s_1,s_2) q(t) q(x_1|s_1,t) q(x_2|s_2,t) q(y|x_1,x_2) \left(2^{h(s_1,s_2)-\imath(y;x_1,x_2)} \right.$$
$$\left. + 2^{h(s_1,s_2|k)-\imath(y;x_1,x_2|k,t)} + 2^{h(s_1|s_2)-\imath(y;x_1|s_2,t,x_2)} + 2^{h(s_2|s_1)-\imath(y;x_2|s_1,t,x_1)} + 1 \right)^{-1}$$
$$= \mathbb{E}_{q_{S_1 S_2 T X_1 X_2 Y_1 Y_2}} \left(1 + 2^{h(S_1|S_2)-\imath(Y;X_1|X_2,S_2,T)} + 2^{h(S_2|S_1)-\imath(Y;X_2|X_1,S_1,T)} \right.$$
$$\left. + 2^{h(S_1,S_2|K)-\imath(Y;X_1,X_2|K,T)} + 2^{h(S_1,S_2)-\imath(Y;X_1,X_2)} \right)^{-1}. \qquad (107)$$

This concludes the proof. ∎

REFERENCES

[1] V. Strassen, Asymptotische Abschätzungen in Shannon's Informations theorie, in Trans. Third. Prague Conf. Inf. Theory, 1962, pp. 689723.

[2] Y. Polyanskiy, H. V. Poor, and S. Verdú, Channel coding in the finite blocklength regime, *IEEE Trans. Inf. Theory*, 56(5), 2307 2359, 2010.

[3] V. Kostina, S. Verdú, "Lossy joint source-channel coding in the finite blocklength regime", arXiv:1209.1317, Sep. 2012.

[4] D. Wang, A. Ingber, and Y. Kochman, The dispersion of joint source- channel coding, in Allerton Conference, 2011, arXiv:1109.6310.

[5] V. Y. F. Tan and O. Kosut, On the dispersions of three network information theory problems, arXiv:1201.3901, Feb 2012.

[6] V. Y. F. Tan, "Transmission of Correlated Sources over a MAC: A Gaussian Approximation-Based Analysis," in Allerton Conference, 2012.

[7] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing ", Physical Review Letters, 2012.

[8] M. Berta, M. Christandl and R. Renner, "The quantum reverse shannon theorem based on one-shot information theory," Commun. Math. Phys. 306, 579615, 2011.

[9] S. Verdú, "Non-Asymptotic Achievability Bounds in Multiuser Information Theory", Allerton Conference, Oct. 2012.

[10] S. Watanabe, S. Kuzuoka, V. Y. F. Tan, "Non-Asymptotic and Second-Order Achievability Bounds for Coding With Side-Information," arXiv:1301.6467, Jan 2013.

[11] P. Cuff, "Distributed channel synthesis," arXiv:1208.4415, Aug. 2012.

[12] A. El Gamal, Y.-H. Kim, "Network Information Theory," Cambridge University Press, 2011.

[13] T. Berger, "Multiterminal Source Coding," Chapter in "The Information Theory Approach to Communications" (G. Longo, editor), Springer-Verlag, 1978.

[14] S.-Y. Tung, "Multiterminal source coding", Ph.d. Thesis, Cornell university, NY, 1978.

[15] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Trans. on Inf. Theory*, 31(6), pp. 727–734, 1985.

[16] A. H. Kaspi, "Rate-distortion function when side-information may be present at the decoder," *IEEE Trans. on Inf. Theory*, 40(6), pp. 2031–2034, 1994.

[17] Z. Zhang and T. Berger, "New results in binary multiple descriptions", *IEEE Trans. on Inf. Theory*, 33(4), pp. 502–521, 1987.

[18] J. Wang, J. Chen, L. Zhao, P. Cuff and H. permuter, "On the role of the refinement layer in multiple description coding and scalable coding", *IEEE Trans. on Inf. Theory*, 57(3), pp. 1443–1456, 2011.

[19] T. Cover, A. El. Gamal and M. Salehi, "Multiple Access Channels with Arbitrarily Correlated Sources," *IEEE Trans. Inf. Theory*, 26(6), 648–657, 1980.

[20] V. Bentkus, "On the dependence of the BerryEsseen bound on dimension," Journal of Statistical Planning and Inference, 113(2), 385-402, 2003.