Successive Refinement with Common Receiver Reconstructions

Badri N. Vellambi* and Roy Timo†

*Institute for Telecommunications Research, University of South Australia
†Communications and Electrical Department, Telecom ParisTech
Email: *badri.vellambi@unisa.edu.au, †timo.roy-clinton@telecom-paristech.fr

***Abstract*—We study the variant of the successive refinement problem where the receivers require identical reconstructions. We characterize the rate region when the joint support of the source and the side information variables is the Cartesian product of their individual supports. The characterization indicates that the side information can be fully used to reduce the communication rates via binning; however, the reconstruction functions can depend only on the Gács-Körner common randomness shared by the two receivers. Unlike existing (inner and outer) bounds to the rate region of the general successive refinement problem, the characterization for the variant studied requires only one auxiliary random variable.**

I. Introduction

We consider the successive refinement (SR) problem where the reconstructions at the receivers are required to be identical. We call this the common receiver reconstructions (CRR) requirement. In this problem, the encoder compresses the source into two messages – one that is common and is intended for both receivers; and the other is private and is intended for only one of the receivers. Each receiver has side information correlated with the source. The receivers use their received messages along with the side information to generate source reconstructions that: (a) meet certain fidelity requirements, and (b) are identical to one another. The problem considered can be viewed as an abstraction of the communication scenario that could arise when conveying data (e.g., meteorological or geological survey data, or an MRI scan) over a network for storage in separate data clusters storing past records of the data. The records, which serve as side information, could be an earlier survey data or a previous scan, depending on the case. The framework considered here arises when data is to be communicated over a degraded broadcast channel using a separate source-channel coding paradigm.

Characterization of rate region of the general successive refinement problem in open and only inner and outer bounds exist [1], [2]. The version of the SR problem where the private message is absent, known as the Heegard-Berger (HB) problem, is open as well [3], [4]. However, complete characterization exists for specific settings of both SR and HB problems. For example, the rate region of the SR problem is known when the side information of the receiver that receives one message is a degraded version of side information of the other receiver [1]. Similarly, the rate region of the HB problem is known when the side information is degraded [3].

The common reconstruction variant of the Wyner-Ziv problem was first motivated and solved by Steinberg [5]. Common reconstructions in other problems were then considered in [6]. In our previous work [7], we characterised the rate region for several cases of the HB problem with CRR requirement. In this work, we characterize the rate region for the SR problem with the CRR requirement (SR-CRR problem) when the joint distribution of the source and side information variables satisfy a certain support condition. The characterization indicates that while the respective side information can be completely used for binning, only the Gács-Körner common randomness between the two side information variables can be used for generating the reconstructions. This feature is also seen in the characterization for the HB problem with the CRR requirement. Unlike the best-known bounds for the SR problem, the characterization of the SR-CRR rate region (when the source satisfies a certain support condition) requires only one auxiliary random variable that is decoded by both receivers. Thus, the CRR requirement obviates the need for a second auxiliary random variable to absorb the private message.

The paper is organized as follows. Section II defines the notation used, and Sec. III defines the problem studied. Section IV presents useful terminologies and their properties, Sec. V presents the main result and its proofs, and the Appendix provides the proofs of the auxiliary results of Sec. IV.

II. Notation

For a set T, and $a, b \in T$, we let $\mathbb{1}(a, b) = 1$ if $a = b$, and 0, otherwise. Uppercase letters (e.g., S, U, V) denote random variables (RVs), and the script versions (e.g., $\mathcal{S}$, $\mathcal{U}$, $\mathcal{V}$) denote their alphabets. All alphabets are assumed to be finite. Realizations of RVs are given by lowercase letters (e.g., x, u, v). Vectors are indicated by superscripts, and their components by subscripts. So, $x^n \triangleq (x_1, \ldots, x_n)$. We denote $x^{n\setminus i} \triangleq (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$. For RVs A, B, C, we denote $A - B - C$ if they form a Markov chain. Given RVs A and B, we denote $A \equiv B$ iff $H(A|B) = H(B|A) = 0$. The probability of an event E is given by $\mathbb{P}(E)$, and $\mathbb{E}$ denotes the expectation operator. $\mathtt{supp}(\cdot)$ denotes the support of a random variable, and $\mathtt{cl}(\cdot)$ denotes the (topological) closure. Lastly, for a p.m.f. p_{AB}, $\sigma_A \triangleq \min\{p_A(a) : p_A(a) > 0\}$ and $\sigma_{A|B} \triangleq \min\{p_{A|B}(a|b) : p_{A|B}(a|b) > 0, ; p_B(b) > 0\}$.

978-1-4799-5186-4/14/$31.00 ©2014 IEEE

III. Problem Definition

The aim of the successive refinement problem is to encode a discrete memoryless source p_S with $S \in \mathcal{S}$ into a pair of messages – one is broadcasted to two receivers, and the other is sent exclusively to one of the receivers. As is given in Fig. 1, the receiver that receives one message has side information $U \in \mathcal{U}$, and the other that receives both the messages has side information $V \in \mathcal{V}$. These side information variables and the source are together generated by a discrete memoryless source with p.m.f. p_{SUV}. This work focuses on *common receiver reconstructions*, where the aim of the encoder is to compress S so that: (1) each receiver can reconstruct the source to within a prescribed level of distortion under a given distortion measure; and (2) the two receiver reconstructions are almost always identical. We aim to characterize the rate region for this problem, which is formally defined as follows.

Definition 1: Given a p.m.f. p_{SUV} and a bounded distortion measure $d : \mathcal{S} \times \hat{\mathcal{S}} \to [0, \bar{D}]$, we say common receiver reconstructions at distortion $D \geq 0$ are *achievable* at a rate pair $(r_{uv}, r_v)^T$, if for each $\varepsilon > 0$, there exist: (a) $n \in \mathbb{N}$, (b) encoders $E_{uv} : \mathcal{S}^n \to \{1, \ldots, \lceil 2^{n(r_{uv}+\varepsilon)} \rceil\}$ and $E_v : \mathcal{S}^n \to \{1, \ldots, \lceil 2^{n(r_v+\varepsilon)} \rceil\}$, and (c) receiver reconstruction functions $D_u : \{1, \ldots, \lceil 2^{n(r_{uv}+\varepsilon)} \rceil\} \times \mathcal{U}^n \to \hat{\mathcal{S}}^n$, and $D_v : \{1, \ldots, \lceil 2^{n(r_{uv}+\varepsilon)} \rceil\} \times \{1, \ldots, \lceil 2^{n(r_v+\varepsilon)} \rceil\} \times \mathcal{V}^n \to \hat{\mathcal{S}}^n$ s.t.

$$\begin{aligned} \hat{S}^n &\triangleq D_u(E_{uv}(S^n), U^n) \\ \tilde{S}^n &\triangleq D_v(E_{uv}(S^n), E_v(S^n), V^n) \end{aligned} \tag{1}$$

satisfy:

$$\begin{aligned} \mathbb{P}[\hat{S}^n \neq \tilde{S}^n] &\leq \varepsilon \\ \textstyle\sum_{i=1}^n \mathbb{E}[d(S_i, \hat{S}_i)] &\leq n(D+\varepsilon) \; . \\ \textstyle\sum_{i=1}^n \mathbb{E}[d(S_i, \tilde{S}_i)] &\leq n(D+\varepsilon) \end{aligned} \tag{2}$$

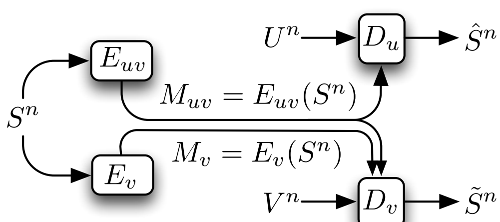

Fig. 1. The successive refinement problem.

We define the rate region $\mathscr{R}_{\mathsf{SR-CRR}}(D)$ for the successive refinement problem with the CRR requirement as the set of all rate pairs achievable in the above sense. Without loss of generality, we may assume that $D \in [\underline{D}, \bar{D}]$, where $\underline{D} \triangleq \min_{\phi: \mathcal{S} \to \hat{\mathcal{S}}} \mathbb{E}[d(S, \phi(S))]$. This minimum distortion $\underline{D}$ is attained when S is conveyed to the receivers.

IV. Additional Terminologies

Given a p.m.f. p_{XY} on $\mathcal{X} \times \mathcal{Y}$, define bipartite graph $\mathbb{G}^{X,Y}[p_{XY}]$ with left nodes $\mathcal{X}$, right nodes $\mathcal{Y}$ and edges between $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ iff $p_{XY}(x, y) > 0$. Define an equivalence relation on $\mathcal{Y}$ by $y_1 \rightleftharpoons y_2$ iff they are in the same connected component of $\mathbb{G}^{X,Y}[p_{XY}]$. Let $\mathsf{GK}^{X,Y} : \mathcal{Y} \to \mathcal{Y}$ satisfy $\mathsf{GK}^{X,Y}(y_1) = \mathsf{GK}^{X,Y}(y_2)$ iff $y_1 \rightleftharpoons y_2$.

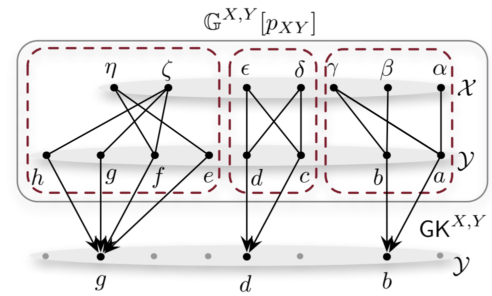

Fig. 2. Ilustration of $\mathbb{G}^{X,Y}[p_{XY}]$ and $\mathsf{GK}^{X,Y}$.

Fig. 2 illustrates an example of $\mathbb{G}^{X,Y}[p_{XY}]$ and a choice for $\mathsf{GK}^{X,Y}$. While there are many choices for $\mathsf{GK}^{X,Y}$, any two are equivalent, i,e., given two choices $\mathsf{GK}_1^{X,Y}$ $\mathsf{GK}_2^{X,Y}$, $\mathsf{GK}_1^{X,Y}(Y) \equiv \mathsf{GK}_2^{X,Y}(Y)$. In this work, we use any one choice to represent $\mathsf{GK}^{X,Y}$ and use the resulting RV $\mathsf{GK}^{X,Y}(Y)$ as the representation of the Gács-Körner common randomness between X and Y [8]. We now define the spaces of p.m.f.'s and associated regions required for the rate region characterization.

Definition 2: Given p_{SUV}, $k \in \mathbb{N}$, $D \geq \underline{D}$ and $\varepsilon > 0$:

1. $\mathcal{P}^{\dagger}_{\varepsilon,D,k}$ denotes the set of p.m.f.'s q_{ABCSUV} s.t. (a) $q_{SUV} = p_{SUV}$, $(A, B, C) - S - (U, V)$, $\max\{|\mathcal{A}|, |\mathcal{B}|, |\mathcal{C}|\} \leq k$; and (b) there exist $f : \mathcal{A} \times \mathcal{C} \times \mathcal{V} \to \hat{\mathcal{S}}$ and $g : \mathcal{A} \times \mathcal{B} \times \mathcal{U} \to \hat{\mathcal{S}}$ s.t.

$$\mathbb{P}[f(A, C, V) \neq g(A, B, U)] \leq \varepsilon \tag{3}$$

$$\mathbb{E}\left[d\big(S, f(A, C, V)\big)\right] \leq D \tag{4}$$

2. $\mathcal{P}^{\ddagger}_{D,k}$ denotes the set of all p.m.f.'s q_{ABCSUV} s.t. (a) $q_{SUV} = p_{SUV}$, $(A, B, C) - S - (U, V)$, $\max\{|\mathcal{A}|, |\mathcal{B}|, |\mathcal{C}|\} \leq k$; and (b) there exists a function $f : \mathcal{A} \times \mathcal{C} \times \mathcal{V} \to \hat{\mathcal{S}}$, such that

$$\mathbb{E}[d(S, f(\mathsf{GK}^{ABU,ACV}(A, C, V)))] \leq D. \tag{5}$$

3. $\mathcal{P}^{*}_{D,k}$ denotes the set of all p.m.f.'s q_{ASUV} s.t. (a) $q_{SUV} = p_{SUV}$, $A - S - (U, V)$, $|\mathcal{A}| \leq k$; and (b) there exists function $f : \mathcal{A} \times \mathcal{V} \to \hat{\mathcal{S}}$, such that

$$\mathbb{E}[d(S, f(\mathsf{GK}^{AU,AV}(A, V)))] \leq D. \tag{6}$$

Definition 3: Given p_{SUV}, $k \in \mathbb{N}$, and $D \geq \underline{D}$, let

$$\mathscr{R}^{\dagger}_{\varepsilon,k}(D) \triangleq \bigcup_{q \in \mathcal{P}^{\dagger}_{\varepsilon,D,k}} \left\{ \begin{pmatrix} r_{uv} \\ r_v \end{pmatrix} : \begin{array}{r} r_{uv} \geq I(S; AB|U) \\ r_v + r_{uv} \geq \begin{pmatrix} I(S; B|ACUV) \\ +I(S; AC|V) \end{pmatrix} \end{array} \right\}.$$

Also, let $\mathscr{R}^{\ddagger}_{k}(D)$ be the resultant region when the union over $q \in \mathcal{P}^{\dagger}_{\varepsilon,D,k}$ above is replaced by a union over $q \in \mathcal{P}^{\ddagger}_{D,k}$. Let

$$\mathscr{R}^{*}_{k}(D) \triangleq \mathtt{cl}\left[\bigcup_{q \in \mathcal{P}^{*}_{D,k}} \left\{ \begin{pmatrix} r_{uv} \\ r_v \end{pmatrix} : \begin{array}{r} r_{uv} \geq I(S; A|U) \\ r_v + r_{uv} \geq I(S; A|V) \end{array} \right\} \right].$$

Remark 1: For any p_{SUV}, $D \geq \underline{D}$, $k \in \mathbb{N}$, and $\varepsilon > 0$,
$\mathcal{P}^{\ddagger}_{D,k} \subseteq \mathcal{P}^{\dagger}_{\varepsilon,D,k}$ and $\mathscr{R}^{*}_{k}(D) \subseteq \mathscr{R}^{\ddagger}_{k}(D) \subseteq \mathscr{R}^{\dagger}_{\varepsilon,k}(D)$. ■

V. The Main Result

The achievablity of $\mathscr{R}^{*}_{|\mathcal{S}|+3}(D)$ follows from a standard proof employing letter typicality, and is omitted. The main result here is the converse when the source p.m.f. satisfies

$$\mathtt{supp}((S, U, V)) = \mathtt{supp}(S) \times \mathtt{supp}((U, V)). \tag{7}$$

Before we prove the converse, four ancillary results quantifying how the three regions of Def 3 are related are presented. Proofs of the first two results are given in the Appendix. The first result is the key step in reducing the symbol error between the two reconstructions from ε to 0, i.e., to prove that for each p.m.f. $q \in \mathcal{P}^{\dagger}_{\varepsilon,D,k}$ satisfying (3) and (4) can be tweaked slightly to a p.m.f. $q' \in \mathcal{P}^{\ddagger}_{D',k}$ for some D' marginally larger than D.

Lemma 1: Let $k \in \mathbb{N}$, $D \geq \underline{D}$, $0 < \varepsilon < \frac{\sigma^4_{SUV}}{10000|\mathcal{S}|^4}$, and p_{SUV} satisfying (7) be given. Let $(r_{uv}, r_v)^T \in \mathscr{R}^{\dagger}_{\varepsilon,k}(D)$ and $\Xi(x) \triangleq \frac{|\mathcal{S}|x}{\sigma_{SUV}-|\mathcal{S}|x} \log \frac{|\mathcal{S}|^2(\sigma_{SUV}-|\mathcal{S}|x)}{x}$. Then,

$$(r_{uv} + \Xi(2\sqrt[4]{\varepsilon}), r_v + 2\,\Xi(2\sqrt[4]{\varepsilon}))^T \in \mathscr{R}^{\ddagger}_{k}\Big(D + \tfrac{2\sqrt{\varepsilon}\bar{D}}{\sigma_S - \sqrt{\varepsilon}}\Big).$$

Lemma 2: Let p_{SUV}, $k \in \mathbb{N}$ and $D \geq \underline{D}$ be given. If p_{SUV} satisfies (7), then $\mathscr{R}^{\ddagger}_{k}(D) \subseteq \mathscr{R}^{*}_{k^2}(D)$.

Lemma 3: For any p.m.f. p_{SUV}, $k \in \mathbb{N}$ and $D \geq \underline{D}$,

$$\mathscr{R}^{*}_{k}(D) \subseteq \mathscr{R}^{*}_{|\mathcal{S}|+3}(D).$$

Proof: The proof is identical to that of Lemma 3 of [7] with the exception of requiring one more function to preserve the value of $I(S; A|U)$. The proof is therefore omitted. ∎

Lemma 4: Let p.m.f. p_{SUV}, sequence of rate pairs $\{(r_{uvi}, r_{vi})^T\}_{i\in\mathbb{N}}$, and sequence $\{D_i\}_{i\in\mathbb{N}}$ be given such that: (a) $\lim_{i\to\infty}(r_{uvi}, r_{vi})^T = (r^*_{uv}, r_v{}^*)^T$; (b) $\lim_{i\to\infty} D_i = D^* > \underline{D}$; and (c) $(r_{uvi}, r_{vi})^T \in \mathscr{R}^{*}_{|\mathcal{S}|+3}(D_i)$.
Then, $(r^*_{uv}, r^*_v)^T \in \mathscr{R}^{*}_{|\mathcal{S}|+3}(D^*)$. Further, if p_{SUV} satisfies (7), the claim holds even if $D^* = \underline{D}$.

Proof: The proof uses time-sharing arguments to show $\mathscr{R}^{*}_{|\mathcal{S}|+3}(D)$ changes smoothly with D, and is omitted. ∎

We are now ready to prove the main result. The proof first uses standard arguments to identify suitable auxiliary RVs and derive an outer bound that does not, *per se*, seem achievable. It is with the use of above lemmas that the outer bound is then reduced to the correct form.

Theorem 1 (Converse): Let p.m.f. p_{SUV} satisfying (7) be given. Let $D \geq \underline{D}$. Then, $\mathscr{R}_{\mathsf{SR-CRR}}(D) \subseteq \mathscr{R}^{*}_{|\mathcal{S}|+3}(D)$.

Proof of Converse: Let $(r_{uv}, r_v)^T \in \mathscr{R}_{\mathsf{SR-CRR}}(D)$. Let $0 < \varepsilon < \frac{\sigma^4_{SUV}}{10000|\mathcal{S}|^4}$. By Def 1, there exist functions E_{uv}, E_v, D_u, D_v satisfying (2). Let $M_{uv} \triangleq E_{uv}(S^n)$, $M_v \triangleq E_v(S^n)$, $\hat{S}^n \triangleq D_u(M_{uv}, U^n)$, and $\tilde{S}^n \triangleq D_v(M_{uv}, M_v, V^n)$. Then,

$$\begin{aligned} n(r_{uv} + \varepsilon) &\geq H(M_{uv}) \geq H(M_{uv}|U^n) \geq I(S^n; M_{uv}|U^n) \\ &= \sum_{i=1}^{n} I(S_i; M_{uv}|U^n S^{i-1}) = \sum_{i=1}^{n} I(S_i; M_{uv} U^{n\setminus i} S^{i-1}|U_i) \\ &\triangleq \sum_{i=1}^{n} I(S_i; A_i B_i|U_i), \end{aligned} \tag{8}$$

where $A_i \triangleq (M_{uv}, U^{i-1})$ and $B_i \triangleq (U^n_{i+1} S^{i-1})$. Similarly,

$$n(r_{uv} + r_v + 2\varepsilon) \geq H(M_{uv} M_v) \geq I(S^n; M_{uv} M_v|V^n) \tag{9}$$

$$= I(S^n; M_{uv} M_v U^n|V^n) - I(S^n; U^n|M_{uv} M_v V^n) \tag{10}$$

$$= \sum_{i=1}^{n}\left[\begin{array}{c} I(S_i; M_{uv} M_v U^n V^{n\setminus i} S^{i-1}|V_i) \\ -I(S^n; U_i|M_{uv} M_v V^n U^{i-1}) \end{array}\right] \tag{11}$$

$$= \sum_{i=1}^{n}\left[\begin{array}{c} I(S_i; A_i B_i M_v V^{n\setminus i} U_i|V_i) - H(U_i|A_i M_v V^n) \\ +H(U_i|A_i M_v V^n S^n) \end{array}\right] \tag{12}$$

$$\overset{(a)}{=} \sum_{i=1}^{n}\left[\begin{array}{c} I(S_i; A_i B_i C_i U_i|V_i) - H(U_i|A_i C_i V_i) \\ +H(U_i|A_i C_i V_i S_i) \end{array}\right] \tag{13}$$

$$= \sum_{i=1}^{n}\left[\, I(S_i; A_i B_i C_i U_i|V_i) - I(S_i; U_i|A_i C_i V_i)\,\right] \tag{14}$$

$$= \sum_{i=1}^{n}\left[\, I(S_i; A_i C_i|V_i) + I(S_i; B_i|A_i C_i U_i V_i)\,\right], \tag{15}$$

where in (a) we have denoted $C_i \triangleq (M_v, V^{n\setminus i})$ and have used $U_i - (M_{uv}, M_v, V^n, U^{i-1}, S_i) - (M_{uv}, M_v, V^n, U^{i-1}, S^n)$.

Now, note that $H(M_{uv} M_v V^n|A_i C_i V_i) = 0$ and $H(M_{uv} U^n|A_i B_i U_i) = 0$. Hence, $\hat{S}_i$ and $\tilde{S}_i$ are functions of (A_i, B_i, U_i) and (A_i, C_i, V_i), resp. Let $\hat{S}_i \triangleq g_i(A_i, B_i, U_i)$ and $\tilde{S}_i \triangleq f_i(A_i, C_i, V_i)$. Define RVs $Q \in \{1, \ldots, n\}$, $\bar{A} \in \{1, \ldots, n\} \times \cup_{i=1}^n \mathcal{A}_i$, $\bar{B} \in \cup_{i=1}^n \mathcal{B}_i$, $\bar{C} \in \cup_{i=1}^n \mathcal{C}_i$, $\bar{S} \in \mathcal{S}$, $U \in \mathcal{U}$ and $V \in \mathcal{V}$ as follows. For $i, j \in \{1, \ldots, n\}$, $a \in \mathcal{A}_i$

$$\begin{aligned} p_{\bar{A},Q}((j,a), i) &= \tfrac{p_{A_i}(a)}{n}\mathbb{1}(i,j) \\ p_{\bar{B}\bar{C}\bar{S}UV|\bar{A}=(i,a),Q=i} &= p_{B_i C_i S_i U_i V_i|A_i=a}, \end{aligned} \tag{16}$$

By construction, we have $p_{\bar{S}UV} = p_{SUV}$ and for $1 \leq i \leq n$, $(A_i, B_i, C_i) - S_i - (U_i, V_i)$ and hence $(\bar{A}, \bar{B}, \bar{C}) - \bar{S} - (U, V)$. Further, the rate expressions (8) and (15) single-letterize to

$$\begin{aligned} r_{uv} + \varepsilon &\geq I(\bar{S}; \bar{A}\bar{B}|U) \\ r_{uv} + r_v + 2\varepsilon &\geq I(\bar{S}; \bar{A}\bar{C}|V) + I(\bar{S}; \bar{B}|\bar{A}\bar{C}UV) \end{aligned} \tag{17}$$

Define maps $\bar{g} : \bar{\mathcal{A}} \times \bar{\mathcal{B}} \times \mathcal{U} \to \hat{\mathcal{S}}$ and $\bar{f} : \bar{\mathcal{A}} \times \bar{\mathcal{C}} \times \mathcal{V} \to \hat{\mathcal{S}}$ by

$$\begin{aligned} \bar{g}((i,a), b, u) &= g_i(a, b, u) \\ \bar{f}((i,a), c, v) &= f_i(a, c, v) \end{aligned} \tag{18}$$

Then,

$$\begin{aligned} \varepsilon &\overset{(b)}{\geq} \sum_{i=1}^{n} \frac{\mathbb{P}[\hat{S}_i \neq \tilde{S}_i]}{n} = \sum_{i=1}^{n} \frac{\mathbb{P}[f_i(A_i, C_i, V_i) \neq g_i(A_i, B_i, U_i)]}{n} \\ &= \mathbb{P}[\bar{f}(\bar{A}, \bar{C}, V) \neq \bar{g}(\bar{A}, \bar{B}, U)], \end{aligned} \tag{19}$$

where (b) follows from (2). Additionally,

$$\mathbb{E}[d(\bar{S}, \bar{f}(\bar{A}, \bar{C}, V)] = \sum_{i=1}^{n} \frac{\mathbb{E}[d(S_i, f_i(A_i, C_i, V_i)]}{n} \overset{(c)}{\leq} D + \varepsilon,$$

where (c) holds because $f_i(A_i, C_i, V_i) = \tilde{S}_i$. Let us denote $\bar{k} \triangleq \max\{|\bar{\mathcal{A}}|, |\bar{\mathcal{B}}|, |\bar{\mathcal{C}}|\}$. Then, by Def. 3,

$$\begin{aligned} &(r_{uv} + \varepsilon, r_v + \varepsilon)^T \in \mathscr{R}^{\dagger}_{\varepsilon,\bar{k}}(D + \varepsilon) \\ &\overset{(a)}{\Rightarrow} \begin{pmatrix} r_{uv} + \varepsilon + \Xi(2\sqrt[4]{\varepsilon}) \\ r_v + \varepsilon + 2\,\Xi(2\sqrt[4]{\varepsilon}) \end{pmatrix} \in \mathscr{R}^{\ddagger}_{\bar{k}}\big(D + \varepsilon + \tfrac{2\sqrt{\varepsilon}\bar{D}}{\sigma_S - \sqrt{\varepsilon}}\big) \\ &\overset{(b)}{\Rightarrow} \begin{pmatrix} r_{uv} + \varepsilon + \Xi(2\sqrt[4]{\varepsilon}) \\ r_v + \varepsilon + 2\,\Xi(2\sqrt[4]{\varepsilon}) \end{pmatrix} \in \mathscr{R}^{*}_{|\mathcal{S}|+3}\big(D + \varepsilon + \tfrac{2\sqrt{\varepsilon}\bar{D}}{\sigma_S - \sqrt{\varepsilon}}\big), \end{aligned}$$

where (a) follows from Lemma 1, (b) from Lemmas 2 and 3. The proof is then complete after invoking Lemma 4. ∎

Appendix

A. Proof of Lemma 1

Let $(r_{uv}, r_u)^T \in \mathscr{R}^{\dagger}_{\varepsilon,k}(D)$ for $\varepsilon < \frac{\sigma^4_{SUV}}{10000|\mathcal{S}|^4}$. Then, there must exist a p.m.f. $p_{ABCSUV} \in \mathcal{P}^{\dagger}_{\varepsilon,D,k}$ such that

$$\begin{aligned} r_{uv} &\geq I(S; AB|U) \\ r_v + r_{uv} &\geq I(S; AC|V) + I(S; B|ACUV). \end{aligned} \tag{20}$$

Let f, g be reconstruction functions meeting (3), (4). Define $\mathcal{E}^* \subseteq \mathcal{A} \times \mathcal{B} \times \mathcal{C}$ by $(a, b, c) \in \mathcal{E}^*$ if

$$\mathbb{P}\big[f(A,C,U) \neq g(A,B,U) | (A,B,C) = (a,b,c)\big] \leq \sqrt{\varepsilon} \quad (21)$$

Then, by Markov inequality and (3), $\mathbb{P}[(A,B,C) \notin \mathcal{E}^*] \leq \sqrt{\varepsilon}$. Note that for $(a,b,c) \in \mathcal{E}^*$, (21) can be rewritten as

$$\sum_{u,v} p_{UV|ABC}(u,v|a,b,c) \mathbb{1}[f(a,c,v), g(a,b,u)] \geq 1 - \sqrt{\varepsilon}.$$

Since p_{SUV} satisfies (7), we have the following.

$$\min_{\substack{s \in \mathrm{supp}(S) \\ (u,v) \in \mathrm{supp}((U,V))}} \left(\frac{p_{UV|S}(u,v|s)}{p_{UV}(u,v)} \right) \geq \sigma_{UV|S} > 0. \quad (22)$$

Then, note that for any $(u,v) \in \mathrm{supp}((U,V))$,

$$p_{UV|ABC}(u,v|a,b,c) = \sum_{s \in \mathrm{supp}(S)} p_{UVS|ABC}(u,v,s|a,b,c)$$
$$\geq \sum_{s \in \mathrm{supp}(S)} \sigma_{UV|S} p_{UV}(u,v) p_{S|ABC}(s|a,b,c) \geq \sigma_{UV|S}\sigma_{UV} \quad (23)$$

Using (23) in (21), we see that for $(a,b,c) \in \mathcal{E}^*$

$$\sum_{(u,v) \in \mathrm{supp}((U,V))} \big(1 - \mathbb{1}[f(a,c,v), g(a,b,u)]\big) \leq \frac{\sqrt{\varepsilon}}{\sigma_{UV}\sigma_{UV|S}} < 1.$$

Hence, it must be the case that when $(a,b,c) \in \mathcal{E}^*$,

$$\mathbb{P}\big[f(A,C,U) \neq g(A,B,U) | (A,B,C) = (a,b,c)\big] = 0 \quad (24)$$

Since the two reconstruction functions agree exactly on $\mathcal{E}^*$, and since $\mathbb{P}[(A,B,C) \in \mathcal{E}^*] \geq 1 - \sqrt{\varepsilon}$, it is possible to prune the distribution p_{ABCS} to have a support $\mathcal{E}^* \times \mathrm{supp}(S)$. To do so, define p.m.f. $p^*_{A^*B^*C^*S}$ over $\mathcal{A} \times \mathcal{B} \times \mathcal{C} \times \mathcal{S}$ by

$$p^*_{A^*B^*C^*S}(a,b,c,s) = \begin{cases} \frac{p_{ABCS}(a,b,c,s)}{\mathbb{P}[(A,B,C) \in \mathcal{E}^* | S = s]}, & \substack{(a,b,c) \in \mathcal{E}^* \\ s \in \mathrm{supp}(S)} \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

Extend $p_{A^*B^*C^*S}$ to $p^*_{A^*B^*C^*SUV}$ s.t. $p^*_{SUV} = p_{SUV}$ and $(A^*, B^*, C^*) - S - (U, V)$. From (24) and (25), we have

$$\mathbb{P}[f(A^*, C^*, U) \neq g(A^*, B^*, V)] = 0 \quad (26)$$

Further, by a simple triangle inequality, we see that

$$\begin{aligned} \mathbb{E}[d(S, f(A^*, C^*, V))] &\leq \mathbb{E}[d(S, f(A,C,V))] \\ &\quad + \bar{D}\, \|p^*_{A^*C^*SV} - p_{ACSV}\|_1 \\ &\overset{(a)}{\leq} D + \bar{D}\Big(\frac{2\sqrt{\varepsilon}}{\sigma_S - \sqrt{\varepsilon}}\Big), \end{aligned}$$

where (a) follows from P1 of Lemma 6. Hence, $p^* \in \mathcal{P}^\dagger_{D',k}$, where $D' \triangleq D + \frac{2\sqrt{\varepsilon}D}{\sigma_S - \sqrt{\varepsilon}}$. Now, from P6 of Lemma 6, we get

$$\begin{aligned} \Delta_1 &\triangleq \big|I(S; A^*B^*|U) - I(S; AB|U)\big| \\ &= \big|H(S|ABU) - H(S|A^*B^*U)\big| \leq \Xi(2\sqrt[4]{\varepsilon}) \end{aligned} \quad (27)$$

Similarly,

$$\begin{aligned} \Delta_2 &\triangleq \big|I(S; A^*C^*|V) - I(S; AC|V)\big| \\ &= \big|H(S|ACV) - H(S|A^*C^*V)\big| \leq \Xi(2\sqrt[4]{\varepsilon}) \end{aligned} . \quad (28)$$

Now, consider $\Delta_3 \triangleq \big|I(S; B^*|A^*C^*UV) - I(S; B|ACUV)\big|$. Expressing Δ_3 using conditional entropies and comparing corresponding terms using P5 and P6 of Lemma 6 yields:

$$\Delta_3 \leq \Xi(\sqrt{\varepsilon}) + \Xi(2\sqrt[4]{\varepsilon}) \leq 2\,\Xi(\sqrt[4]{\varepsilon}).$$

Combining the bounds for $\Delta_1, \Delta_2, \Delta_3$ with (20), we have

$$\begin{aligned} r_{uv} + \Xi(2\sqrt[4]{\varepsilon}) &\geq I(S; A^*B^*|U) \\ r_v + r_{uv} + 3\,\Xi(2\sqrt[4]{\varepsilon}) &\geq I(S; B^*|A^*C^*UV) + I(S; A^*C^*|V). \end{aligned}$$

Hence, the claim follows. ■

B. Proof of Lemma 2

Pick $p_{ABCSUV} \in \mathcal{P}^\ddagger_{D,k}$. Let $f : \mathcal{A} \times \mathcal{C} \times \mathcal{V} \to \hat{\mathcal{S}}$ satisfy

$$\mathbb{E}[d(S, f(\mathsf{GK}^{ABU,ACV}(A,C,V)))] \leq D. \quad (29)$$

An application of Lemma 5 with $A_1 = (A, C)$ and $A_2 = (A, B)$ yields the following conclusion.

$$\mathsf{GK}^{ACV,ABU}(A,C,V) \equiv (\mathsf{GK}^{AC,AB}(A,C), \mathsf{GK}^{V,U}(V)) \quad (30)$$

Now, define $\tilde{A} \triangleq \mathsf{GK}^{AB,AC}(A,C)$ and let $\tilde{p}_{\tilde{A}SUV}$ denote the p.m.f. of $(\tilde{A}, S, U, V)$. Then, by construction the Markov chain $\tilde{A} - S - (U, V)$ holds. Further,

$$H(\tilde{A}, \mathsf{GK}^{U,V}(V) | \mathsf{GK}^{\tilde{A}V,\tilde{A}U}(\tilde{A}, V)) = 0 \quad (31)$$

Then, by (30), $H(\mathsf{GK}^{ABU,ACV}(A,C,V)|\mathsf{GK}^{\tilde{A}V,\tilde{A}U}(\tilde{A},V)) = 0$. Let $\tilde{f} : \tilde{\mathcal{A}} \times \mathcal{V} \to \mathcal{A} \times \mathcal{C} \times \mathcal{V}$ be such that

$$\tilde{f}(\mathsf{GK}^{\tilde{A}U,\tilde{A}V}(\tilde{A}, V)) = \mathsf{GK}^{ABU,ACV}(A,C,V). \quad (32)$$

Then, from (29), $\mathbb{E}[d(S, f(\tilde{f}(\mathsf{GK}^{\tilde{A}V,\tilde{A}U}(\tilde{A}, V))))] \leq D$. Hence, $\tilde{p}_{\tilde{A}SUV} \in \mathcal{P}^*_{D,k^2}$ since $\tilde{\mathcal{A}} \subseteq \mathcal{A} \times \mathcal{C}$. Lastly, note that

$$\begin{aligned} r_{uv} &\geq I(S; AB|U) \geq I(S; \mathsf{GK}^{AC,AB}(A,C)|U) = I(S; \tilde{A}|U) \\ r_v + r_{uv} &\geq I(S; AC|V) \geq I(S; \mathsf{GK}^{AC,AB}(A,C)|V) = I(S; \tilde{A}|V) \end{aligned}$$

The claim holds since p_{ABCSUV} was chosen at random. ■

C. Ancillary Results

Lemma 5: Let p.m.f. $p_{A_1A_2SUV}$ over $\mathcal{A}_1 \times \mathcal{A}_2 \times \mathcal{S} \times \mathcal{U} \times \mathcal{V}$ be s.t.: (a) $(A_1, A_2) - S - (U, V)$, and (b) (7) is met. Then,

$$\mathsf{GK}^{A_1V,A_2U}(A_1, V) \equiv (\mathsf{GK}^{A_1,A_2}(A_1), \mathsf{GK}^{U,V}(V)). \quad (33)$$

Proof: Let $\eta^* \triangleq \min_{(s,u,v) \in \mathrm{supp}(S,U,V)} \frac{p_{UV|S}(u,v|s)}{p_{UV}(u,v)}$. Then, $\eta^* > 0$ due to (7) . Now, let $(a_1, a_2) \in \mathrm{supp}(A_1, A_2)$ and $(u, v) \in \mathrm{supp}(U, V)$. Then,

$$\begin{aligned} p_{A_1VA_2U}(a_1, v, a_2, u) &= \sum_{s \in \mathrm{supp}(S)} p_{A_1A_2|S}(a_1, a_2, s) p_{UV|S}(u, v|s) \\ &\geq \sum_{s \in \mathrm{supp}(S)} \eta^* p_{A_1A_2S}(a_1, a_2, s) p_{UV}(u, v) \\ &= \eta^* p_{A_1A_2}(a_1, a_2) p_{UV}(u, v) > 0. \end{aligned}$$

Hence, for $(a_1, v), (a_1', v') \in \mathcal{A}_1 \times \mathcal{V}$, $(a_1, v) \rightleftharpoons (a_1', v')$ in $\mathbb{G}^{A_1V,A_2U}[p_{A_1VA_2U}]$ if and only if $a_1 \rightleftharpoons a_1'$ in $\mathbb{G}^{A_1,A_2}[p_{A_1A_2}]$ and $v \rightleftharpoons v'$ in $\mathbb{G}^{U,V}[p_{UV}]$. Hence, the claim follows. ■

Lemma 6: Let p.m.f. $p_{A_1A_2SB_1B_2}$ be given such that $(A_1,A_2) - S - (B_1,B_2)$ and (7) is met. Let $\delta < \sigma_{SB_1B_2}$ and $\mathcal{E} \subseteq \mathcal{A}_1 \times \mathcal{A}_2$ be such that $\mathbb{P}[(A_1,A_2) \in \mathcal{E}] \geq 1-\delta$. Define p.m.f. $\tilde{p}_{\tilde{A}_1\tilde{A}_2S}$ by

$$\tilde{p}_{A_1A_2S}(a_1,a_2,s) = \begin{cases} \frac{p_{A_1A_2S}(a_1,a_2,s)}{1-\mathbb{P}[(A_1,A_2)\notin\mathcal{E}|S=s]}, & \substack{(a_1,a_2)\in\mathcal{E} \\ s\in\text{supp}(S)} . \\ 0, & \text{otherwise} \end{cases}$$

and extend it to a p.m.f. $\tilde{p}_{\tilde{A}_1\tilde{A}_2SB_1B_2}$ by setting $(\tilde{A}_1,\tilde{A}_2) - S - (B_1,B_2)$ and $\tilde{p}_{SB_1B_2} = p_{SB_1B_2}$. Let $\Xi(\cdot)$ be as defined in Lemma 1. Then the following hold.

P1 $\left\|\tilde{p}_{\tilde{A}_1\tilde{A}_2SB_1B_2} - p_{A_1A_2SB_1B_2}\right\|_1 \leq \frac{2\delta}{\sigma_S-\delta}$.

P2 If $(a_1a_2,b_2,b_2)\in\text{supp}((A_1,A_2,B_1,B_2))$ and $(a_1,a_2)\in\mathcal{E}$,

$$\left\|\begin{matrix}\tilde{p}_{S|\tilde{A}_2\tilde{A}_2B_1B_2}(\cdot|a_1,a_2,b_1,b_2) \\ -p_{S|A_1A_2B_1B_2}(\cdot|a_1,a_2,b_1,b_2)\end{matrix}\right\|_1 \leq \frac{\delta}{\sigma_S-\delta}$$

P3. Let $\mathcal{D} = \left\{a_1 \in \mathcal{A}_1 : \mathbb{P}[(A_1,A_2) \notin \mathcal{E}|A_1=a_1] \leq \sqrt{\delta}\right\}$. Then, $\mathbb{P}[A_1 \in \mathcal{D}] \geq 1-\sqrt{\delta}$.

P4. If $a_1 \in \mathcal{D}$, and $b_1 \in \mathcal{B}_1$ s.t. $p_{B_1A_1}(b_1,a_1) > 0$, then

$$\left\|p_{S|\tilde{A}_1B_1}(\cdot|a_1,b_1) - p_{S|A_1B_1}(\cdot|a_1,b_1)\right\|_1 \leq \frac{2\sqrt{\delta}|\mathcal{S}|}{\sigma_{SB_1B_2}-2\sqrt{\delta}|\mathcal{S}|}.$$

Additionally, if $\delta \leq \frac{\sigma^2_{SB_1B_2}}{100|\mathcal{S}|^2}$, then the following hold:

P5. $|H(S|\tilde{A}_1\tilde{A}_2B_1B_2) - H(S|A_1A_2B_1B_2)| \leq \Xi(\delta)$.

P6. $|H(S|\tilde{A}_1B_1) - H(S|A_1B_1)| \leq \Xi(2\sqrt{\delta})$.

Proof: To prove P1, first note that the contribution to the L_1-norm due to elements of $\mathcal{E}^c \times \mathcal{S} \times \mathcal{B}_1 \times \mathcal{B}_2$ equals $\mathbb{P}[(A_1,A_2) \notin \mathcal{E}] \leq \delta \leq \frac{\delta}{\sigma_S-\delta}$. When $(a_1,a_2) \in \mathcal{E}$ and $(a_1,a_2,s,b_1,b_2) \in \text{supp}((A_1,A_2,S,B_1,B_2))$, we see that

$$1 \leq \frac{\tilde{p}_{\tilde{A}_1\tilde{A}_2SB_1B_2}(a_1,a_2,s,b_1,b_2)}{p_{A_1A_2SB_1B_2}(a_1,a_2,s,b_1,b_2)} \leq \frac{\sigma_S}{\sigma_S-\delta}, \tag{34}$$

since $\mathbb{P}[(A_1,A_2) \notin \mathcal{E}|S=s] = \frac{\mathbb{P}[(A_1,A_2)\in\mathcal{E},S=s]}{p_S(s)} \leq \frac{\delta}{\sigma_S}$. Rearranging and summing (34) over all elements of the set $\mathcal{E} \times \mathcal{S} \times \mathcal{B}_1 \times \mathcal{B}_2$ yields an upper bound of $\frac{\delta}{\sigma_S-\delta}$ for the contribution from $\mathcal{E} \times \mathcal{S} \times \mathcal{B}_1 \times \mathcal{B}_2$.

P2 follows from straightfoward algebraic manipulations of (34), and is omitted. We now prove the validity of P3.

$$\delta \geq \mathbb{P}[(A_1,A_2)\notin\mathcal{E}] \geq \sum_{a_1\in\mathcal{D}^c} p_{A_1}(a_1)\mathbb{P}[(A_1,A_2)\notin\mathcal{E}|A_1=a_1]$$
$$\geq \sqrt{\delta}\cdot\mathbb{P}[A_1\in\mathcal{D}^c] = \sqrt{\delta}\left(1-\mathbb{P}[A_1\in\mathcal{D}]\right). \tag{35}$$

To prove P4, we use the following arguments.

$$\tilde{p}_{S|\tilde{A}_1B_1}(s|a_1,b_1) = \frac{\sum_{a_2}\tilde{p}_{S\tilde{A}_1\tilde{A}_2B_1}(s,a_1,a_2,b_1)}{\sum\limits_{a_2,s'}\tilde{p}_{S\tilde{A}_1\tilde{A}_2B_1}(s',a_1,a_2,b_1)} \tag{36}$$

$$= \frac{\sum\limits_{a_2:(a_1,a_2)\in\mathcal{E}} \frac{p_{SA_1A_2B_1}(s,a_1,a_2,b_1)}{1-\mathbb{P}[(A_1,A_2)\notin\mathcal{E}|S=s]}}{\sum_{s'\in\mathcal{S},\,a_2:(a_1,a_2)\in\mathcal{E}} \frac{p_{SA_1A_2B_1}(s',a_1,a_2,b_1)}{1-\mathbb{P}[(A_1,A_2)\notin\mathcal{E}|S=s']}} \tag{37}$$

$$\leq \frac{\left(\frac{\sigma_S}{\sigma_S-\delta}\right)p_{SA_1B_1}(s,a_1,b_1)}{p_{A1,B1}(a_1,b_1) - \mathbb{P}[(A_1,A_2)\notin\mathcal{E},A_1=a_1]} \tag{38}$$

$$= \frac{p_{S|A_1B_1}(s|a_1,b_1)}{1-\frac{\sqrt{\delta}}{p_{B_1|A_1}(b_1|a_1)}} \cdot \frac{1}{1-\frac{\delta}{\sigma_S}} \overset{(a)}{\leq} \frac{p_{S|A_1B_1}(s|a_1,b_1)}{1-\frac{\sqrt{\delta}}{\sigma_{B_1|S}}-\frac{\delta}{\sigma_S}}, \tag{39}$$

where in (a) we have used

$$\min_{b_1,a_1} p_{B_1|A_1}(b_1|a_1) = \min_{b_1,a_1}\sum_s p_{S|A_1}(s)p_{B_1|S}(b_1|s) \geq \sigma_{B_1|S}.$$

Using (37), we see that

$$\tilde{p}_{S|\tilde{A}_1B_1}(s|a_1,b_1) \geq \frac{\sum\limits_{a_2:(a_1,a_2)\in\mathcal{E}} p_{SA_1A_2B_1}(s,a_1,a_2,b_1)}{\left(\frac{\sigma_S}{\sigma_S-\delta}\right)p_{A_1B_1}(a_1,b_1)}$$
$$= \frac{p_{SA_1B_1}(s,a_1,b_1) - \sum\limits_{a_2:(a_1,a_2)\notin\mathcal{E}} p_{SA_1A_2B_1}(s,a_1,a_2,b_1)}{\left(\frac{\sigma_S}{\sigma_S-\delta}\right)p_{A_1B_1}(a_1,b_1)}$$
$$\geq \left(1-\tfrac{\delta}{\sigma_S}\right)\left(p_{S|A_1B_1}(s|a_1b_1) - \tfrac{\sqrt{\delta}}{p_{B_1|A_1}(b_1|a_1)}\right) \tag{40}$$
$$\geq p_{S|A_1B_1}(s|a_1b_1) - \tfrac{\delta}{\sigma_S} - \tfrac{\sqrt{\delta}}{\sigma_{B_1|S}}. \tag{41}$$

A straighforward argument combining (39) and (41) along with the fact that $\frac{\delta}{\sigma_S} + \frac{\sqrt{\delta}}{\sigma_{B_1|S}} \leq \frac{2\sqrt{\delta}}{\sigma_{SB_1B_2}}$ yields

$$\left\|\tilde{p}_{S|\tilde{A}_1B_1}(\cdot|a_1b_1) - p_{S|A_1B_1}(\cdot|a_1b_1)\right\|_1 \leq \tfrac{2\sqrt{\delta}|\mathcal{S}|}{\sigma_{SB_1B_2}-2\sqrt{\delta}|\mathcal{S}|}.$$

P5 follows from P2 upon an application of Lemma 2.7 of [9]. The proof of P5 is omitted since it is similar to that of P6 outlined below. Since $\delta < \frac{\sigma^2_{SB_1B_2}}{100|\mathcal{S}|^2}$, Lemma 2.7 of [9] guarantees that when $a_1 \in \mathcal{D}$, and $p_{B_1A_1}(b_1,a_1) > 0$,

$$\left|\begin{matrix}H(S|A_1{=}a_1,B_1{=}b_1) \\ -H(S|\tilde{A}_1{=}a_1,B_1{=}b_1)\end{matrix}\right| \leq \tfrac{2\sqrt{\delta}|\mathcal{S}|}{\sigma_{SB_1B_2}-2\sqrt{\delta}|\mathcal{S}|}\log\tfrac{\sigma_{SB_1B_2}-2\sqrt{\delta}|\mathcal{S}|}{2\sqrt{\delta}}.$$

Further, when $a_1 \notin \mathcal{D}$, the difference in the conditional entropies can be bounded by $\log|\mathcal{S}|$. Combining these two conditional entropy bounds with P1 and P3, we obtain P6. ■

ACKNOWLEDGEMENT

This research was supported under the Australian Research Council's Discovery Project DP120102123.

REFERENCES

[1] Y. Steinberg and N. Merhav, "On Successive refinement for the Wyner-Ziv Problem," *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1636–1654, 2004.

[2] C. Tian and S. Diggavi, "Side-information scalable source coding," *IEEE Trans. Inform. Theory*, vol. 54, no. 12, pp. 5591–5607, 2008.

[3] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Trans. Inform. Theory*, vol. 31, no. 6, pp. 727–734, 1985.

[4] A. Kaspi, "Rate-distortion function when side-information may be present at the decoder," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 2031–2034, 1994.

[5] Y. Steinberg, "Coding and common reconstruction," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 4995–5010, 2009.

[6] B. Ahmadi, R. Tandon, O. Simeone, and H. V. Poor, "Heegard-Berger and cascade source coding problems with common reconstruction constraints," *IEEE Trans. Inform. Theory*, vol. 59, no. 3, pp. 1458–1474, 2013.

[7] B. N. Vellambi and R. Timo, "The heegard-berger problem with common receiver reconstructions," in *Proc. 2013 Inform. Theory Workshop (ITW 2013)*, Seville, Spain pp. 1–5, 2013.

[8] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems of Control and Inform. Theory*, vol. 2, no. 2, pp. 149–162, 1973.

[9] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Orlando, USA: Academic Press, 1982.