Utility of Encoder Side Information for the Lossless Kaspi/Heegard-Berger Problem

Thomas Laich
Signal & Inform. Proc. Laboratory
ETH Zurich
Email: tlaich@student.ethz.ch

Michèle Wigger
Comm. and Electr. Department
Telecom ParisTech
Email: michele.wigger@telecom-paristech.fr

***Abstract*—We consider the lossless Kaspi/Heegard-Berger source coding problem where an encoder communicates a common description of two sources to two decoders, and each decoder wants to reconstruct one of the sources with the help of side information. We present new results on the utility of encoder side information for this scenario. We show that for some sources and side informations—e.g., for some instances of conditionally less noisy side information—the minimum rate that is required to describe the sources is strictly reduced when the side information is also known at the encoder. On the other hand, we identify classes of sources and side informations—e.g., physically degraded side information—where encoder side information does not change the minimum description rate.**

We show similar results for a scenario where one decoder has to reconstruct both sources and for a scenario where the encoder is informed only about one of the decoder's side information.

I. Introduction

Consider the lossless Kaspi/Heegard-Berger source-coding problem [1], [2] (Figure 1 without dashed arrows) in which an encoder observes two correlated, memoryless sources $X_1^n \triangleq (X_{1,1},\dots,X_{1,n})$ and $X_2^n \triangleq (X_{2,1},\dots,X_{2,n})$ and communicates a common description M to Decoders 1 and 2. Decoder 1 observes the side information $Y_1^n \triangleq (Y_{1,1},\dots,Y_{1,n})$ and wishes to reconstruct X_1^n losslessly based on Y_1^n and M. Similarly, Decoder 2 wishes to reconstruct X_2^n losslessly based on $Y_2^n \triangleq (Y_{2,1},\dots,Y_{2,n})$ and M. We refer to this setup as the scenario with an *ignorant encoder* for the encoder does not observe the side informations Y_1^n and Y_2^n.

The minimum rate of M that is required to describe the two sources losslessly in this scenario, R^*_{ign}, is known for the following sources and side informations:

(a) *physically degraded* side information, i.e., $(X_1^n, X_2^n) \to Y_1^n \to Y_2^n$ forms a Markov chain [2];
(b) Y_1^n is *conditionally less noisy* [3] than Y_2^n given X_2^n and $H(X_2^n \,|\, Y_1^n) \le H(X_2^n \,|\, Y_2^n)$;
(c) equal sources $X_1^n = X_2^n$ [4]; and
(d) complementary side information $Y_1^n = X_2^n$ and $Y_2^n = X_1^n$ [4].

Notice that case (b) includes case (a) as a special case.

The main interest of this paper is in the slightly modified scenario with an *informed encoder*, in which the encoder knows the side informations Y_1^n and Y_2^n (see Figure 1 with dashed arrows). We are particularly interested in answering the question whether the minimum description rate with informed

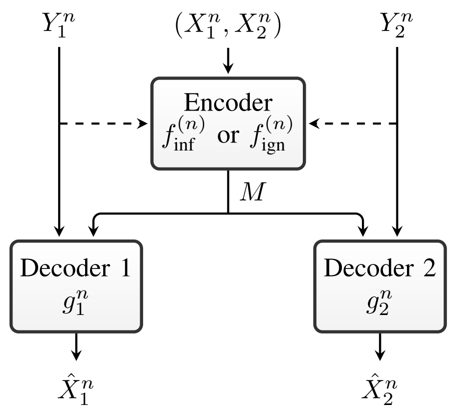

Fig. 1. The Kaspi/Heegard-Berger lossless source-coding problem with two sources. With dashed arrows it depicts the scenario with an informed encoder and without the dashed arrows the scenario with an ignorant encoder.

encoder, R^*_{inf}, can be strictly smaller than the minimum description rate with ignorant encoder, i.e.,

$$R^*_{\text{inf}} < R^*_{\text{ign}}. \tag{1}$$

In fact, up to present, R^*_{inf}, is known only in the above cases (c) and (d), where

$$R^*_{\text{inf}} = R^*_{\text{ign}}. \tag{2}$$

Thus, in particular, when the two decoders are interested in reconstructing the same source $X_1^n = X_2^n$, then encoder side information does not reduce the minimum description rate.

This is different for the *lossy* version of our source coding problem, where encoder side information can strictly reduce the minimum description rate [1], [5], [6].

In this paper, we show that this is also the case when the two decoders wish to losslessly reconstruct two different sources. That means, we show that in our problem (1) holds for certain sources and side informations, e.g., for some instances of conditionally less noisy side information. We further show that for some sources and side informations the minimum description rate is reduced even when the encoder knows only one of the decoder's side information, e.g., Y_1^n but not Y_2^n.

On the other hand, we prove that for a certain class of sources and side informations knowledge of both side informations Y_1^n and Y_2^n at the encoder does not reduce the minimum description rate in our lossless problem and thus (2) holds.

This class includes physically degraded side information as special case.

We also present single-letter characterizations of the minimum description rate with informed encoder for some classes of sources and side informations.

Lastly, we consider a modified scenario, in which one of the decoders wishes to reconstruct both sources. We show that also here encoder side information can strictly reduce the minimum description rate, and we characterize the minimum description rates (both with informed and with ignorant encoder) for some classes of sources and side informations.

II. Problem Statement

The setup is characterized by four finite alphabets $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_1$, and $\mathcal{Y}_2$ and a joint probability law $P_{X_1X_2Y_1Y_2}$ over these alphabets. Let the tuple (X_1, X_2, Y_1, Y_2) be distributed according to $P_{X_1X_2Y_1Y_2}$, and let $(X_1^n, X_2^n, Y_1^n, Y_2^n)$ denote n independently and identically distributed copies of this tuple.

In the informed-encoder scenario, the encoder produces a common description

$$M \triangleq f_{\text{inf}}^{(n)}(X_1^n, X_2^n, Y_1^n, Y_2^n)$$

using some encoder mapping $f_{\text{inf}}^{(n)} : \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{Y}_1^n \times \mathcal{Y}_2^n \to \mathcal{M}$. In the ignorant-encoder scenario, it produces

$$M \triangleq f_{\text{ign}}^{(n)}(X_1^n, X_2^n)$$

using some mapping $f_{\text{ign}}^{(n)} : \mathcal{X}_1^n \times \mathcal{X}_2^n \to \mathcal{M}$. In both scenarios, each decoder $j \in \{1, 2\}$ produces its reconstruction sequence

$$\hat{X}_j^n \triangleq g_j^{(n)}(M, Y_j^n)$$

using some decoder mapping $g_j^{(n)} : \mathcal{M} \times \mathcal{Y}_j^n \to \mathcal{X}_j^n$.

A triple $(f_{\text{inf}}^{(n)}, g_1^{(n)}, g_2^{(n)})$ or $(f_{\text{ign}}^{(n)}, g_1^{(n)}, g_2^{(n)})$ consisting of an encoder and two decoder mappings is called an *n-block source code* with informed or with ignorant encoder, respectively. Its *description rate* is defined as

$$\kappa^{(n)} \triangleq \frac{1}{n} \log_2 |\mathcal{M}|$$

and the *average joint error probability* as

$$P_e^{(n)} \triangleq \Pr[\hat{X}_1^n \neq X_1^n \text{ or } \hat{X}_2^n \neq X_2^n].$$

A rate $R \geq 0$ is said to be *achievable* with informed encoder if for every $\epsilon > 0$ and all sufficiently large integers n there exists an n-block source code $(f_{\text{inf}}^{(n)}, g_1^{(n)}, g_2^{(n)})$ such that $P_e^{(n)} < \epsilon$ and $\kappa^{(n)} \leq R + \epsilon$. In the scenario with an ignorant encoder, achievability is defined similarly. We define the *minimum description rate*:

$$R_{\text{inf}}^* \triangleq \min\{R \geq 0 : R \text{ achievable with informed encoder}\},$$
$$R_{\text{ign}}^* \triangleq \min\{R \geq 0 : R \text{ achievable with ignorant encoder}\}.$$

Since the encoder can always ignore the side information,

$$R_{\text{inf}}^* \leq R_{\text{ign}}^*. \tag{3}$$

Note that R_{ign}^* depends on the joint distribution $P_{X_1X_2Y_1Y_2}$ only via the marginal distributions $P_{X_1X_2Y_1}$ and $P_{X_1X_2Y_2}$; the same is not true for R_{inf}^*.

Remark 1. The scenario with an informed encoder can be viewed as a scenario with an ignorant encoder if in the latter the sources are augmented with the side informations, i.e., if X_1 is replaced by (X_1, Y_1) and X_2 by (X_2, Y_2).

The following special cases will be considered. We say that the side information is *physically degraded* if

$$(X_1, X_2) \to Y_1 \to Y_2 \quad \text{or} \quad (X_1, X_2) \to Y_2 \to Y_1$$

forms a Markov chain. The side information Y_1 is *conditionally less noisy* than Y_2 given the source component X_2 (abbreviated $(Y_1 \succeq Y_2 \,|\, X_2)$) [3], if for all finite random variables W that form the Markov chain $W \to (X_1, X_2) \to (Y_1, Y_2)$:

$$I(W; Y_1 \,|\, X_2) \geq I(W; Y_2 \,|\, X_2).$$

We use the notation $(X_1, X_2) \sim \text{DSBS}(p)$ to indicate that (X_1, X_2) is a *doubly-symmetric binary source with parameter* p [7]. Similarly, $E \sim \text{Bern}(p)$ indicates that E is a Bernoulli-p random variable. We further use the notation $\text{BSC}(p)$ for a *binary symmetric channel* with parameter p and $\text{BEC}(q)$ for a *binary erasure channel* with parameter q [7].

III. Bounds on R_{IGN}^* and R_{INF}^*

Lemma 1 (Theorem 2 in [8])**.** *With an ignorant encoder:*

$$R_{ign}^* \leq \min_W \Big\{ \max\{I(W; X_1X_2 \,|\, Y_1), I(W; X_1X_2 \,|\, Y_2)\} + H(X_1 \,|\, WY_1) + H(X_2 \,|\, WY_2) \Big\}$$

where the minimization is over all finite random variables W that take value in a set $\mathcal{W}$ of size $|\mathcal{W}| \leq |\mathcal{X}_1 \times \mathcal{X}_2| + 3$ and that satisfy the Markov chain $W \to (X_1, X_2) \to (Y_1, Y_2)$.

Proposition 1. *With an informed encoder:*

$$R_{inf}^* \geq \max \Big\{ H(X_1 \,|\, Y_1) + H(X_2 \,|\, X_1Y_1Y_2), H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1Y_2) \Big\}$$

and

$$R_{inf}^* \leq \min_W \Big\{ \max\{I(W; X_1X_2Y_2 \,|\, Y_1), I(W; X_1X_2Y_1 \,|\, Y_2)\} + H(X_1 \,|\, WY_1) + H(X_2 \,|\, WY_2) \Big\}$$

where the minimization is over all finite random variables W that take value in a set $\mathcal{W}$ of size $|\mathcal{W}| \leq |\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}_1 \times \mathcal{Y}_2| + 3$.

Proof: The lower bound follows by revealing the pair (Y_1, Y_2) to one of the decoders and by Lemma 2-(a) on the next page. The upper bound follows by evaluating Lemma 1 for the augmented sources (X_1, Y_1) and (X_2, Y_2), see Remark 1. ■

Remark 2. The upper and lower bounds in Proposition 1 do not coincide in general (see [9, Proposition 3.9]).

IV. EXACT RESULTS FOR R^*_{IGN} AND R^*_{INF}

The following existing results for ignorant encoder will serve for comparison.

Lemma 2 (Previous results for ignorant encoder)**.** *The minimum description rate R^*_{ign} is known for the following distributions $P_{X_1X_2Y_1Y_2}$:*

(a) If the side information is physically degraded*, i.e., $(X_1, X_2) \to Y_1 \to Y_2$ forms a Markov chain, then [2]*

$$R^*_{ign} = H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1).$$

(b) If $(Y_1 \succeq Y_2 \,|\, X_2)$ and $H(X_2 \,|\, Y_1) \le H(X_2 \,|\, Y_2)$, then [3]

$$R^*_{ign} = H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1).$$

(c) If we have equal sources $X_1 = X_2 = X$*, then [4]*

$$R^*_{ign} = \max_{j\in\{1,2\}} H(X \,|\, Y_j). \tag{4}$$

(d) If the decoders have complementary side information*, i.e., $Y_1 = X_2$ and $Y_2 = X_1$, then [4]*

$$R^*_{ign} = \max\{H(X_1 \,|\, X_2), H(X_2 \,|\, X_1)\}.$$

Further results on R^*_{ign} are obtained by exchanging the indices 1 and 2 in the results above.

Remark 3. In cases (c) and (d) the minimum description rate for an informed encoder coincides with the minimum description rate for an ignorant encoder, i.e.,

$$R^*_{\text{inf}} = R^*_{\text{ign}}. \tag{5}$$

In case (c), equality (5) follows from (4) and because $\max_{j\in\{1,2\}} H(X_j \,|\, Y_j)$ is a lower bound on the minimum description rate even in the scenario with informed encoder. In case (d), equality (5) holds because the sources (X_1, X_2) determine the side information Y_1 and Y_2.

Theorem 1 (Informed encoder)**.** *For the following classes of distributions $P_{X_1X_2Y_1Y_2}$, we can identify the minimum description rate with informed encoder R^*_{inf}:*

(i) If the side information is physically degraded*, i.e., $(X_1, X_2) \to Y_1 \to Y_2$ forms a Markov chain, then*

$$R^*_{inf} = H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1).$$

(ii) If $X_1 \to (X_2, Y_1) \to Y_2$ forms a Markov chain and $H(X_2 \,|\, Y_1) \le H(X_2 \,|\, Y_2)$, then

$$R^*_{inf} = H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1).$$

(iii) If $X_2 \to (X_1, Y_1) \to Y_2$ and $X_1 \to (X_2, Y_2) \to Y_1$ form Markov chains, then

$$R^*_{inf} = \max_{j\in\{1,2\}} H(X_1X_2 \,|\, Y_j).$$

(iv) If $H(X_2Y_2 \,|\, Y_1) \le H(X_2 \,|\, Y_2)$, then

$$R^*_{inf} = H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1Y_2).$$

Further results on R^*_{inf} are obtained by exchanging the indices 1 and 2 in the results above.

Proof: By Proposition 1. Specifically, the achievability results follow by specializing the upper bound in Proposition 1 to the choices $W = X_2$ (cases i and ii), $W = (X_1, X_2)$ (case iii), and $W = (X_2, Y_2)$ (case iv). ■

Case (i) is included in case (ii) because $(X_1, X_2) \to Y_1 \to Y_2$ implies that the two conditions in (ii) are satisfied.

Remark 4. The two sets of distributions $P_{X_1X_2Y_1Y_2}$ covered by Theorem 1 and Lemma 2, resp., are not subsets of each other. There are thus setups that we can solve with informed encoder but not with ignorant encoder, and vice versa.

The following three examples illustrate some of the results in Lemma 2 and Theorem 1. The first example satisfies the conditions in (i) but not the ones in (iii) or (iv).

Example 1. Let $(X_1, X_2) \sim \text{DSBS}\,(p)$, $p \in (0, 1/2)$, let $Y_1 = X_1 \oplus X_2$, and let Y_2 be the result of passing Y_1 through a BEC (q), $q \in (0, 1)$. The setup is physically degraded because $(X_1, X_2) \to Y_1 \to Y_2$ forms a Markov chain. Hence, by Lemma 2-(a) and by Theorem 1-(i),

$$R^*_{\text{inf}} = R^*_{\text{ign}} = H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1) = 1.$$

Our second example satisfies the conditions in (ii) but not the conditions in (i), (iii), or (iv).

Example 2 (Example 2 in [3]). Let $(X_1, X_2) \sim \text{DSBS}\,(1/3)$, and let the side information Y_1 and Y_2 be the results of passing X_2 through independent BEC $(2/3)$ and BSC $(1/4)$, respectively. It holds that $X_1 \to (X_2, Y_1) \to Y_2$ forms a Markov chain, and further inspection reveals $(Y_1 \succeq Y_2 \,|\, X_2)$. Moreover,

$$H(X_2 \,|\, Y_1) = 2/3,$$
$$H(X_2 \,|\, Y_2) = H_\text{b}(1/4) \approx 0.81,$$

and hence, $H(X_2 \,|\, Y_1) \le H(X_2 \,|\, Y_2)$. We can apply Lemma 2-(b) and Theorem 1-(ii) to obtain

$$R^*_{\text{inf}} = R^*_{\text{ign}} = H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1)$$
$$= H_\text{b}(1/4) + H_\text{b}(1/3).$$

Our third example satisfies the conditions in (iii) but not the conditions in (i), (ii), or (iv).

Example 3. Let $(X_1, X_2) \sim \text{DSBS}\,(p)$, $p \in (0, 1/2)$. Also, let Y_1 be the result of passing X_2 through a BEC (q_1) and Y_2 be the result of passing X_1 through a BEC (q_2). We assume that the erasure events at the two BECs are independent and that $q_1, q_2 \in (0, 1/2)$. By Theorem 1-(iii),

$$R^*_{\text{inf}} = \max_{j\in\{1,2\}} H(X_1X_2 \,|\, Y_j) = \max\{q_1, q_2\} + H_\text{b}(p).$$

The rate $\max_{j\in\{1,2\}} H(X_1X_2 \,|\, Y_j)$ is also achievable without encoder side information, and thus, also here, $R^*_{\text{ign}} = R^*_{\text{inf}}$.

The next example meets the conditions in case (b) of Lemma 2. But only if $q(2 + H_\text{b}(p)) \le 1$, it also meets the conditions in one of the cases of Theorem 1; namely case (iv).

Proposition 2. *Let $(X_1, X_2) \sim \text{DSBS}\,(p)$, and independent thereof let (E_1, E_2) be a pair of correlated binary random* *variables with $E_j \sim \mathrm{Bern}(q_j)$, for $q_j \in (0,1)$ and $j \in \{1,2\}$. Let $Y_1 \triangleq (\tilde{Y}_1, E_2)$ and $Y_2 \triangleq (\tilde{Y}_2, E_1)$, where*

$$\tilde{Y}_j \triangleq \begin{cases} e & \text{if } E_j = 1, \\ (X_1, X_2) & \text{if } E_j = 0, \end{cases} \qquad j \in \{1,2\}.$$

Then,

$$R^*_{\mathit{inf}} = \max\{q_1, q_2\} + P_{E_1E_2}(1,1) H_{\mathrm{b}}(p), \tag{6a}$$

$$R^*_{\mathit{ign}} = \max\{q_1, q_2\} + \min\{q_1, q_2\} H_{\mathrm{b}}(p). \tag{6b}$$

Proof: Equality (6b) follows from Lemma 2-(b). Equality (6a) follows from Proposition 1 where for the upper bound we choose $W \triangleq (W_1, W_2)$ with

$$W_j \triangleq \begin{cases} X_j & \text{if } E_j = 1, \\ e & \text{if } E_j = 0. \end{cases} \qquad j \in \{1,2\}. \qquad \blacksquare$$

Theorem 2 (Utility of encoder side information)**.**

- *For distributions $P_{X_1X_2Y_1Y_2}$ that satisfy the conditions in cases* (i)*,* (ii)*, or* (iii) *of Theorem 1, encoder side information does not reduce the minimum description rate, i.e.,*

$$R^*_{\mathit{inf}} = R^*_{\mathit{ign}}. \tag{7}$$

- *For some distributions $P_{X_1X_2Y_1Y_2}$, encoder side information strictly reduces the minimum description rate; i.e.,*

$$R^*_{\mathit{inf}} < R^*_{\mathit{ign}}. \tag{8}$$

  This is in particular the case for some distributions that satisfy the conditions in case (iv) *of Theorem 1 or the conditions in case* (b) *of Lemma 2.*

Proof: For case (i), equality (7) follows from Lemma 2-(a) and Theorem 1-(i). For case (ii), equality (7) follows from Lemma 2-(b), Theorem 1-(ii), and the fact that the Markov chain $X_1 \to (X_2, Y_1) \to Y_2$ implies $(Y_1 \succeq Y_2 \,|\, X_2)$. For case (iii), equality (7) follows from Theorem 1-(iii) and by noting that $\max_{j\in\{1,2\}} H(X_1X_2|Y_j)$ is achievable also with an ignorant encoder; in fact, it is achievable by describing the pair (X_1, X_2) to both decoders, see Lemma 2-(c).

On the other hand, if in Proposition 2 we have $P_{E_1E_2}(1,1) = P_{E_1E_2}(0,0) = 0$ and $P_{E_1E_2}(1,0) = q$ and $P_{E_1E_2}(0,1) = 1-q$, for $q \in (0, 1/2)$, then by (6)

$$R^*_{\mathrm{inf}} = 1 - q,$$
$$R^*_{\mathrm{ign}} = 1 - q + qH_{\mathrm{b}}(p),$$

and (8) holds for all $p \in (0,1)$. $\blacksquare$

V. Special Case: Degraded Source Sets

Consider the related scenario where Decoder 1 wants to reconstruct the source pair (X_1, X_2), whereas Decoder 2 is satisfied with reconstructing only the second source X_2. More specifically, our scenario here is as described in Section II but where X_1 has to be replaced by (X_1, X_2) and $\hat{X}_1$ has to be replaced by $(\hat{X}_1, \hat{X}_2^{(1)})$, where $\hat{X}_2^{(1)}$ denotes Decoder 1's reconstruction of the source component X_2.

Theorem 3 (Ignorant encoder)**.** *For degraded source sets and ignorant encoder we have the following results:*

(A) If $H(X_1X_2 \,|\, Y_1) \geq H(X_1X_2 \,|\, Y_2)$ or if $H(X_2 \,|\, Y_1) \geq H(X_2 \,|\, Y_2)$, then

$$R^*_{\mathit{ign}} = H(X_1X_2 \,|\, Y_1).$$

(B) If $(Y_1 \succeq Y_2 \,|\, X_2)$, then

$$R^*_{\mathit{ign}} = \max_{j\in\{1,2\}} H(X_2 \,|\, Y_j) + H(X_1 \,|\, X_2Y_1).$$

(C) If $X_1 \to (X_2, Y_2) \to Y_1$ forms a Markov chain, then

$$R^*_{\mathit{ign}} = \max_{j\in\{1,2\}} H(X_1X_2 \,|\, Y_j).$$

Proof: Replacing in the lower bound of Proposition 1 and in the upper bound of Lemma 1 the source X_1 by the pair (X_1, X_2), results in lower and upper bounds for this setup with degraded source sets and uninformed encoder. The resulting bounds suffice to establish cases (A) and (C). To establish the desired result in case (B), we also need Lemma 2-(b), where again we replace X_1 by (X_1, X_2). Details omitted. $\blacksquare$

Theorem 4 (Informed encoder)**.** *For degraded source sets and informed encoder, we have the following results:*

(I) If $H(X_1X_2 \,|\, Y_1) \geq H(X_1X_2 \,|\, Y_2)$ or if $H(X_2 \,|\, Y_1) \geq H(X_2 \,|\, Y_2)$, then

$$R^*_{\mathit{inf}} = H(X_1X_2 \,|\, Y_1).$$

(II) If $X_1 \to (X_2, Y_1) \to Y_2$ forms a Markov chain, then

$$R^*_{\mathit{inf}} = \max_{j\in\{1,2\}} H(X_2 \,|\, Y_j) + H(X_1 \,|\, X_2Y_1).$$

(III) If $X_1 \to (X_2, Y_2) \to Y_1$ forms a Markov chain, then

$$R^*_{\mathit{inf}} = \max_{j\in\{1,2\}} H(X_1X_2 \,|\, Y_j).$$

(IV) If $H(X_2Y_2 \,|\, Y_1) \leq H(X_2 \,|\, Y_2)$, then

$$R^*_{\mathit{inf}} = H(X_2 \,|\, Y_2) + H(X_1 \,|\, X_2Y_1Y_2).$$

Proof: Follows from the bounds that result when in Proposition 1 the source X_1 is replaced by (X_1, X_2). Details omitted. $\blacksquare$

Remark 5. The set of distributions $P_{X_1X_2Y_1Y_2}$ covered by Lemma 2 is included in the set of distributions covered by Theorem 3. Similarly, the set of distributions $P_{X_1X_2Y_1Y_2}$ covered by Theorem 1 is included in the set of distributions covered by Theorem 4. The following Example 4 illustrates that the inclusions are strict. Thus, with degraded source sets we can derive more results than in the original setup, both for ignorant and for informed encoder.

Example 4. Let $(X_1, X_2) \sim \mathrm{DSBS}(p)$, $p \in (0, 1/2)$, and let (E_1, E_2) be independent of (X_1, X_2) with $E_j \sim \mathrm{Bern}(q_j)$, for $q_j \in (0,1)$ and $j \in \{1,2\}$. Also, define

$$Y_j = \begin{cases} X_j & \text{if } E_j = 0 \\ e & \text{if } E_j = 1. \end{cases}$$

Notice that $X_1 \to (X_2, Y_1) \to Y_2$ forms a Markov chain and $(Y_1 \succ Y_2 | X_2)$. Thus, by Theorem 3-(B) and Theorem 4-(II),

$$\begin{aligned} R^*_{\text{inf}} = R^*_{\text{ign}} &= \max_{j\in\{1,2\}} H(X_2 \,|\, Y_j) + H(X_1 \,|\, X_2 Y_1) \\ &= \max\left\{q_1 + H_{\text{b}}(p), q_2 + q_1 H_{\text{b}}(p)\right\}. \end{aligned}$$

For $q_1 = q_2$, this example does not satisfy the conditions in any of the cases (a)–(d) of Lemma 2 nor the conditions in any of the cases (i)–(iv) of Theorem 1.

The next example is covered by Theorem 3 but not by Theorem 4. In fact, it meets the conditions in case (B) of Theorem 3 for arbitrary $q \in (0, 1/2)$, but the condition in case (IV) of Theorem 4 only when $q \le (2 + H_{\text{b}}(p))^{-1}$.

Proposition 3. *For degraded source sets and the distributions $P_{X_1X_2Y_1Y_2}$ of Proposition 2:*

$$R^*_{\textit{inf}} = \max\left\{q_1(1 + H_{\text{b}}(p)), q_2 + P_{E_1E_2}(1,1)H_{\text{b}}(p)\right\}, \tag{9a}$$
$$R^*_{\textit{ign}} = \max\{q_1, q_2\} + q_1 H_{\text{b}}(p). \tag{9b}$$

Proof: Omitted. Similar to the proof of Proposition 2. ■

In the example of Propositions 2 and 3, the requirement that Decoder 1 also decodes source X_2 increases the minimum description rates R^*_{inf} and R^*_{ign} whenever Decoder 1 gets the worse side information, i.e., whenever $q_1 > q_2$.

Theorem 5 (Utility of encoder side information)**.** *For degraded source sets, we have the following results:*

- *For distributions $P_{X_1X_2Y_1Y_2}$ that satisfy the conditions in cases* (I)–(III) *of Theorem 4, encoder side information does not reduce the minimum description rate, i.e.,*
$$R^*_{\textit{inf}} = R^*_{\textit{ign}}. \tag{10}$$
- *For certain distributions $P_{X_1X_2Y_1Y_2}$, encoder side information reduces the minimum description rate, i.e.,*
$$R^*_{\textit{inf}} < R^*_{\textit{ign}}. \tag{11}$$
This is in particular the case for some distributions that satisfy the conditions in case (IV) *of Theorem 4 or the conditions in case* (B) *of Theorem 3.*

Proof: The first statement follows from Theorems 3 and 4. The second statement follows from Proposition 3 with the same choice of parameters as in the proof of Theorem 2. More precisely, if in Proposition 3 we have $P_{E_1E_2}(1,1) = P_{E_1E_2}(0,0) = 0$, $P_{E_1E_2}(1,0) = q$, and $P_{E_1E_2}(0,1) = 1 - q$, and $q \in (0, 1/2)$, then by (9),

$$\begin{aligned} R^*_{\text{inf}} &= \max\{q(1 + H_{\text{b}}(p)), (1 - q)\}, \\ R^*_{\text{ign}} &= 1 - q + qH_{\text{b}}(p), \end{aligned}$$

and (11) holds for all $p \in (0, 1)$. ■

VI. Extension: Partially Informed Encoder

We reconsider the original setup in Section II but now assume that the encoder knows only the side information Y_1^n but not Y_2^n. The encoder thus produces the common description

$$M \triangleq f^{(n)}_{\text{part}}(X_1^n, X_2^n, Y_1^n)$$

using some mapping $f^{(n)}_{\text{part}} : \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{Y}_1^n \to \mathcal{M}$. The minimum description rate, R^*_{part}, is defined as R^*_{inf} in Section II but where $f^{(n)}_{\text{inf}}$ needs to be replaced by $f^{(n)}_{\text{part}}$. Obviously,

$$R^*_{\text{inf}} \le R^*_{\text{part}} \le R^*_{\text{ign}}.$$

Proposition 4. *Let $(X_1, X_2) \sim \mathrm{DSBS}\,(p)$ and let (E_1, E_2) be independent of (X_1, X_2) with $E_j \sim \mathrm{Bern}\,(q_j)$, for $q_j \in (0, 1)$ and $j \in \{1, 2\}$. Also, let $Y_1 \triangleq \tilde{Y}_1$ and $Y_2 \triangleq (\tilde{Y}_2, E_1)$, where*

$$\tilde{Y}_j \triangleq \begin{cases} e & \text{if } E_j = 1, \\ X_1 & \text{if } E_j = 0, \end{cases} \qquad j \in \{1, 2\}.$$

Then,

$$\begin{aligned} R^*_{\textit{part}} &= \max\left\{q_1 + H_{\text{b}}(p), q_2 + \left(1 - q_2 + P_{E_1E_2}(1,1)\right)H_{\text{b}}(p)\right\}, \\ R^*_{\textit{ign}} &= \max\left\{q_1 + H_{\text{b}}(p), q_2 + (1 - q_2 + q_1)H_{\text{b}}(p)\right\}. \end{aligned}$$

Thus, when $P_{E_1E_2}(1,1) < q_1$, then

$$R^*_{\textit{part}} < R^*_{\textit{ign}}. \tag{12}$$

Notice that for this example the encoder cannot determine Y_2^n. Still, $R^*_{\textit{inf}} = R^*_{\textit{part}}$, i.e., here a partially informed encoder is as good as a fully informed encoder.

Proof: Omitted. Similar to the proof of Proposition 2. ■

We immediately obtain the following theorem:

Theorem 6 (Utility of *partial* encoder side information)**.** *There are distributions $P_{X_1X_2Y_1Y_2}$ for which* (12) *holds and the missing side information at the encoder, Y_2, is not a deterministic function of X_1, X_2, Y_1, i.e, $H(Y_2|X_1, X_2, Y_2) > 0$.*

Acknowledgment

The authors thank the city of Paris for supporting their research under the "Emergences" program. The work of T. Laich was also supported by an IDEA League student grant.

References

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

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

[3] R. Timo, T. Oechering, and M. Wigger, "Source coding problems with conditionally less noisy side information," Dec. 2012, submitted to *IEEE Trans. Inf. Theory*. [Online]. Available: http://arxiv.org/abs/1212.2396

[4] A. Sgarro, "Source coding with side information at several decoders," *IEEE Trans. Inf. Theory*, vol. 23, no. 2, pp. 179–182, 2006.

[5] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, 1976.

[6] S. Diggavi, E. Perron, and E. Telatar, "Lossy source coding with gaussian or erased side-information," in *Proceedings ISIT 2009*, Seoul, Korea, June 28–July 3 2009, pp. 1035–1039.

[7] A. El Gamal and Y.-H. Kim, *Network Information Theory*, 1st ed. Cambridge University Press, 2011.

[8] R. Timo, A. Grant, T. Chan, and G. Kramer, "Source coding for a simple network with receiver side information," in *Proceedings ISIT 2008*, Toronto, Canada, July 6–11 2008, pp. 2307–2311.

[9] T. Laich, "The Kaspi/Heegard-Berger problem with an informed encoder," Master's Thesis, ETH Zürich, Switzerland, 2012.