

# On the Heegard-Berger Problem with Common Reconstruction Constraints

Behzad Ahmadi  
CWCSPR, ECE Dept.  
NJIT

Ravi Tandon  
Dept. of Electrical Engineering  
Princeton University

Oswaldo Simeone  
CWCSPR, ECE Dept.  
NJIT

H. Vincent Poor  
Dept. of Electrical Engineering  
Princeton University

**Abstract**—In lossy source coding with side information at the decoder (i.e., the Wyner-Ziv problem), the estimate of the source obtained at the decoder cannot be generally reproduced at the encoder, due to its dependence on the side information. In some applications this may be undesirable, and a Common Reconstruction (CR) requirement, whereby one imposes that encoder and decoder be able to agree on the decoder’s estimate, may be instead in order. The rate-distortion function under the CR constraint has been recently derived for the point-to-point (Wyner-Ziv) problem. In this paper, this result is extended to the Heegard-Berger (HB) problem and to its variant with cooperating decoders. Specifically, for the HB problem, the rate-distortion function is derived under the assumption that the side information sequences at the two decoders are stochastically degraded. The rate-distortion function is also calculated explicitly for the special case of binary source and erased side information with Hamming distortion metric. The rate-distortion function is then characterized also for the HB problem with cooperating decoders and physically degraded side information.

## I. INTRODUCTION

In source coding problems with side information at a single decoder, such as in the Wyner-Ziv problem, or at multiple decoders, such as in the Heegard-Berger (HB) problem [1], the side information sequences are, in general, used in two different ways. The first is as a means to reduce the rate required for communication between encoder and decoders via binning. The second is, instead, as an additional observation that the decoders can leverage, along with the bits received from the encoder, in order to improve their local estimates. Leveraging the side information in this latter way, while advantageous in terms of rate-distortion trade-off, may have unacceptable consequences for some applications. In fact, this use of the side information entails that the reconstruction  $\hat{X}$  at a decoder cannot be reproduced at the encoder. In applications such as transmission of sensitive medical information, this may not be desirable. Instead, one may want to add the constraint that the reconstruction at the decoder be reproducible by the encoder [2].

This idea, referred to as the Common Reconstruction (CR) constraint, was first proposed in [2], where it is shown that for the Wyner-Ziv problem the rate-distortion function under the CR constraint is given by

$$R_{X|Y}^{CR}(D) = \min I(X; \hat{X}|Y), \quad (1)$$

where the minimum is taken over all probability mass functions (pmfs)  $p(\hat{x}|x)$  such that  $E[d(X, \hat{X})] \leq D$ . Comparing

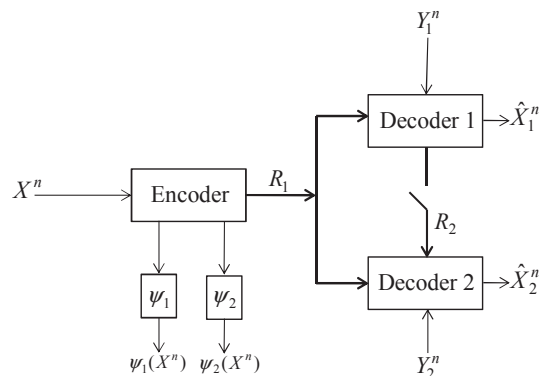


Figure 1. Heegard-Berger source coding problem with common reconstruction (switch open). Closing the switch enables decoder cooperation.

(1) with the conventional Wyner-Ziv rate-distortion function, namely  $R_{X|Y}^{WZ}(D) = \min I(X; U|Y)$  with minimum overall all pmfs  $p(u|x)$  and deterministic functions  $\hat{x}(u, y)$  such that  $E[d(X, \hat{x}(U, Y))] \leq D$ , it can be seen that the additional CR constraint prevents the decoder from using the side information as a means to improve its estimate  $\hat{X}$ .

In this paper, we extend the characterization of the rate-distortion function under the CR constraint of [2] from the point-to-point setting to the HB problem [1] (Fig. 1, switch open), and its variant, studied in [3], in which decoder cooperation is enabled by a limited capacity link from one decoder, Decoder 1, to the other, Decoder 2 (Fig. 1, switch closed). We recall that [1] derived the rate-distortion function for the HB problem (without CR constraints) under the assumption that the side information sequences are stochastically degraded versions of the source  $X^n$ . Instead, for the case with decoder cooperation, inner and outer bounds on the rate distortion region for this problem are obtained in [3] (without CR constraints) under the assumption that the side information of Decoder 2 is physically degraded with respect to that of Decoder 1.

For the HB problem with the CR constraint, in Sec. II, we derive the rate-distortion function under the assumption that the side information sequences are stochastically degraded. We also calculate this function explicitly for binary source and erased side information with the Hamming distortion metric. Moreover, for the HB problem with the CR constraint and decoder cooperation, we derive the rate-distortion region

under the assumption that the side information sequences are (physically) degraded in either direction (Sec. III-A and Sec. III-B). We emphasize that the corresponding problem without the CR constraint is still open as per the discussion above.

## II. HEEGARD-BERGER PROBLEM WITH COMMON RECONSTRUCTION

In this section, we first detail the system model for the HB source coding problem in Fig. 1 (switch open) with CR in Sec. II-A. Next, the characterization of the corresponding rate-distortion performance is derived under the assumption that one of the two side information sequences is a stochastically degraded version of the other in the sense of [1]. Finally, an example with binary sources is worked out in Sec. II-C. The case of Gaussian sources under quadratic distortion is tackled in [4]. The proofs of the results presented in the paper are omitted due to space limitations and can be found in [4].

### A. System Model

In this section the system model for the HB problem with CR is detailed. The system is defined by the pmf  $p_{XY_1Y_2}(x, y_1, y_2)$  and discrete alphabets  $\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, \hat{\mathcal{X}}_1$ , and  $\hat{\mathcal{X}}_2$  as follows. The source sequence  $X^n$  and side information sequences  $Y_1^n$  and  $Y_2^n$ , with  $X^n \in \mathcal{X}^n$ ,  $Y_1^n \in \mathcal{Y}_1^n$ , and  $Y_2^n \in \mathcal{Y}_2^n$  are such that the tuples  $(X_i, Y_{1i}, Y_{2i})$  for  $i \in [1, n]$  are independent and identically distributed (i.i.d.) with joint pmf  $p_{XY_1Y_2}(x, y_1, y_2)$ . The encoder observes a sequence  $X^n$  and encodes it into a message  $J$  of  $nR$  bits, which is delivered to the decoders. Decoders 1 and 2 wish to reconstruct the source sequence  $X^n$  within given distortion requirements, to be discussed below, as  $\hat{X}_1^n \in \hat{\mathcal{X}}_1^n$  and  $\hat{X}_2^n \in \hat{\mathcal{X}}_2^n$ , respectively. The estimated sequence  $\hat{X}_j^n$  is obtained as a function of the message  $J$  and the side information sequence  $Y_j^n$  for  $j = 1, 2$ . The estimates are constrained to satisfy distortion constraints defined by per-symbol distortion metrics  $d_j(x, \hat{x}_j) : \mathcal{X} \times \hat{\mathcal{X}}_j \rightarrow [0, D_{max}]$  with  $0 < D_{max} < \infty$ . The reconstructions  $\hat{X}_2^n$  and  $\hat{X}_1^n$  are also required to satisfy the CR constraints, as formalized below.

**Definition 1.** An  $(n, R, D_1, D_2, \epsilon)$  code for the HB problem with CR consists of an encoding function

$$g: \mathcal{X}^n \rightarrow [1, 2^{nR}], \quad (2)$$

which maps the source sequence  $X^n$  into a message  $J$ ; a decoding function for Decoder 1,

$$h_1: [1, 2^{nR}] \times \mathcal{Y}_1^n \rightarrow \hat{\mathcal{X}}_1^n, \quad (3)$$

which maps the message  $J$  and the side information  $Y_1^n$  into the estimated sequence  $\hat{X}_1^n$ ; a decoding function for Decoder 2

$$h_2: [1, 2^{nR}] \times \mathcal{Y}_2^n \rightarrow \hat{\mathcal{X}}_2^n \quad (4)$$

which maps message  $J$  and the side information  $Y_2^n$  into the estimated sequence  $\hat{X}_2^n$ ; and two reconstruction functions

$$\psi_1: \mathcal{X}^n \rightarrow \hat{\mathcal{X}}_1^n \quad (5a)$$

$$\text{and } \psi_2: \mathcal{X}^n \rightarrow \hat{\mathcal{X}}_2^n, \quad (5b)$$

which map the source sequence into the estimated sequences at the encoder, namely  $\psi_1(X^n)$  and  $\psi_2(X^n)$ , respectively; such that the distortion constraints are satisfied, i.e.,

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E} [d_j(X, \hat{X}_j)] \leq D_j \text{ for } j = 1, 2, \quad (6)$$

and the CR requirements hold, namely,

$$\Pr [\psi_j(X^n) \neq h_j(g(X^n), Y_j^n)] \leq \epsilon, \quad j = 1, 2. \quad (7)$$

Given distortion pairs  $(D_1, D_2)$ , a rate pair  $R$  is said to be achievable if, for any  $\epsilon > 0$  and sufficiently large  $n$ , there exists an  $(n, R, D_1 + \epsilon, D_2 + \epsilon, \epsilon)$  code. The rate-distortion function  $R(D_1, D_2)$  is defined as  $R(D_1, D_2) = \inf\{R : \text{the triple } (R, D_1, D_2) \text{ is achievable}\}$ .

### B. Rate-Distortion Function

In this section, a single-letter characterization of the rate-distortion function for the HB problem with CR is derived, under the assumption that the joint pmf  $p(x, y_1, y_2)$  is such that there exists a conditional pmf  $\tilde{p}(y_1|y_2)$  for which

$$p(x, y_1) = \sum_{y_2 \in \mathcal{Y}_2} p(x, y_2) \tilde{p}(y_1|y_2). \quad (8)$$

In other words, the side information  $Y_1$  is a stochastically degraded version of  $Y_2$ .

**Proposition 2.** *If the side information  $Y_1$  is stochastically degraded with respect to  $Y_2$ , the rate-distortion function for the HB problem with CR is given by*

$$\begin{aligned} R_{HB}^{CR}(D_1, D_2) &= \min I(X; \hat{X}_1|Y_1) + I(X; \hat{X}_2|Y_2\hat{X}_1) \\ &= \min I(X; \hat{X}_1\hat{X}_2|Y_2) + I(\hat{X}_1; Y_2|Y_1), \end{aligned} \quad (9a)$$

where the mutual information terms are evaluated with respect to the joint pmf

$$p(x, y_1, y_2, \hat{x}_1, \hat{x}_2) = p(x, y_1, y_2)p(\hat{x}_1, \hat{x}_2|x), \quad (10)$$

and minimization is performed with respect to the conditional pmf  $p(\hat{x}_1, \hat{x}_2|x)$  under the constraints

$$\mathbb{E}[d_j(X, \hat{X}_j)] \leq D_j, \text{ for } j = 1, 2. \quad (11)$$

The proof of the converse can be found in [4]. Achievability follows as a special case of Theorem 3 of [1] and can be easily shown using standard arguments. In particular, the encoder randomly generates a standard lossy source code  $\hat{X}_1^n$  for the source  $X^n$  with rate  $I(X; \hat{X}_1)$  bits per source symbol. Random binning is used to reduce the rate to  $I(X; \hat{X}_1|Y_1)$ . By the Wyner-Ziv theorem, this guarantees that both Decoder 1 and Decoder 2 are able to recover  $\hat{X}_1^n$  (since  $Y_1$  is a degraded version of  $Y_2$ ). The encoder then maps the source  $X^n$  into the reconstruction sequence  $\hat{X}_2^n$  using a codebook that is generated conditional on  $\hat{X}_1^n$  with rate  $I(X; \hat{X}_2|\hat{X}_1)$  bits per source symbol. Random binning is again used to reduce the rate to  $I(X; \hat{X}_2|Y_2\hat{X}_1)$ . From the Wyner-Ziv theorem, and the fact that Decoder 2 knows the sequence  $\hat{X}_1^n$ , it follows that Decoder 2 can recover the reconstruction  $\hat{X}_2^n$  as well.

Note that, since the reconstruction sequences  $\hat{X}_1^n$  and  $\hat{X}_2^n$  are generated by the encoder, functions  $\psi_1$  and  $\psi_2$  that guarantee the CR constraints (7) exist by construction.

*Remark 3.* If we remove the CR constraint, then the rate-distortion function under the assumption of Proposition 2 is given by [1]

$$R_{HB}(D_1, D_2) = \min I(X; U_1|Y_1) + I(X; U_2|Y_2U_1), \quad (12)$$

where the mutual information terms are evaluated with respect to the joint pmf

$$p(x, y_1, y_2, u_1, u_2) = p(x, y_1, y_2)p(u_1, u_2|x), \quad (13)$$

and minimization is performed with respect to the conditional pmf  $p(u_1, u_2|x)$  and the deterministic functions  $\hat{x}_j(u_j, y_j)$ , for  $j = 1, 2$ , such that distortion constraints  $E[d_j(X, \hat{x}_j(U_j, Y_j))] \leq D_j$ , for  $j = 1, 2$  are satisfied. Comparison of (9) with (12) reveals that, similar to the discussion for the point-to-point set-up, the CR constraint permits the use of side information only to reduce the rate via binning, but not to improve the decoder's estimates via the use of the auxiliary codebooks represented by variables  $U_1$  and  $U_2$ , and functions  $\hat{x}_j(u_j, y_j)$ , for  $j = 1, 2$ , in (13).

*Remark 4.* Consider the case in which the side information sequences are available in a causal fashion in the sense of [5]; that is, the decoding functions (3)-(4) are modified as  $h_{ji}: [1, 2^{nR}] \times \mathcal{Y}_j^i \rightarrow \hat{\mathcal{X}}_{ji}$ , for  $i \in [1, n]$  and  $j = 1, 2$ , respectively. Following similar steps as in [5], it can be concluded that the rate-distortion function in this case is the same as if the two side information sequences were not available at the decoders, and is thus given by (9) upon removing the conditioning on the side information. Note that this is true irrespective of the joint pmf  $p(x, y_1, y_2)$  and hence it holds also for non-degraded side information. This result can be explained by noting that, as explained in [5], causal side information prevents the possibility of reducing the rate via binning. Since the CR constraint also prevents the side information from being used to improve the decoders' estimates, it follows that the side information is useless in terms of rate-distortion performance, if used causally.

### C. Binary Source with Erased Side Information and Hamming or Erasure Distortion

In this section, we consider a binary source  $X \sim \text{Ber}(\frac{1}{2})$  with erased side information sequences  $Y_1$  and  $Y_2$ . The source  $Y_2$  is an erased version of the source  $X$  with erasure probability  $p_2$  and  $Y_1$  is an erased version of  $X$  with erasure probability  $p_1 > p_2$ . This means that  $Y_j = e$ , where  $e$  represents an erasure, with probability  $p_j$  and  $Y_j = X$  with probability  $1 - p_j$ . Note that, with these assumptions, the side information  $Y_1$  is stochastically degraded with respect to  $Y_2$ . In fact, we have the factorization (8), where additional distributions  $p(y_2|y_1)$  and  $\tilde{p}(y_1|y_2)$  are illustrated in Fig. 2. As seen in Fig. 2, the pmf  $\tilde{p}(y_1|y_2)$  is characterized by the probability  $\tilde{p}_1$  that satisfies the equality  $\tilde{p}_1 = p_2 + \tilde{p}_1(1 - p_2)$ . We focus on Hamming and erasure distortions. For the Hamming distortion,

the reconstruction alphabets are binary,  $\hat{\mathcal{X}}_1 = \hat{\mathcal{X}}_2 = \{0, 1\}$ , and we have  $d_j(x, \hat{x}_j) = 0$  if  $x = \hat{x}_j$  and  $d_j(x, \hat{x}_j) = 1$  otherwise for  $j = 1, 2$ . Instead, for the erasure distortion the reconstruction alphabets are  $\hat{\mathcal{X}}_1 = \hat{\mathcal{X}}_2 = \{0, 1, e\}$ , and we have for  $j = 1, 2$ :  $d_j(x, \hat{x}_j) = 0$  if  $\hat{x}_j = x$ ,  $d_j(x, \hat{x}_j) = 1$  if  $\hat{x}_j = e$  and  $d_j(x, \hat{x}_j) = \infty$  otherwise.

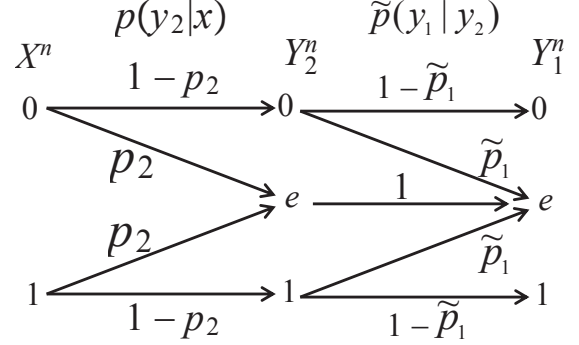


Figure 2. Illustration of the pmfs in the factorization (8) of the joint distribution  $p(x, y_1, y_2)$  for a binary source  $X$  and erased side information sequences  $(Y_1, Y_2)$ .

In [4, Appendix C], it is proved that for the point-to-point set-up with  $X \sim \text{Ber}(\frac{1}{2})$  and erased side information  $Y$  with erasure probability  $p$ , the rate-distortion function with CR under Hamming distortion is given by

$$R_{X|Y}^{CR}(D) = \begin{cases} R_B^{CR}(D, p) \triangleq p(1 - H(D)) & \text{for } D \leq 1/2 \\ 0 & \text{for } D > 1/2, \end{cases} \quad (14)$$

while under erasure distortion we get

$$R_{X|Y}^{CR}(D) = R_{BE}^{CR}(D, p) \triangleq p(1 - D), \quad (15)$$

where  $H(x)$  denotes the binary entropy function. Note that the rate-distortion function with erased side information and Hamming distortion without the CR constraint is derived in [7].

**Proposition 5.** *The rate-distortion function for the HB problem with CR for the binary source with the stochastically degraded erased side information sequences illustrated in Fig. 2 under Hamming distortion is given by*

$$R_{HB}^{CR}(D_1, D_2) = \begin{cases} 0 & \text{if } D_1 \geq 1/2 \text{ and } D_2 \geq 1/2 \\ R_B^{CR}(D_1, p_1) & \text{if } D_1 \leq 1/2 \text{ and } D_2 \geq \min(D_1, 1/2) \\ R_B^{CR}(D_2, p_2) & \text{if } D_1 \geq 1/2 \text{ and } D_2 \leq 1/2 \\ \tilde{R}_{HB}^{CR}(D_1, D_2) & \text{if } D_2 \leq D_1 \leq 1/2 \end{cases} \quad (16)$$

where  $R_B^{CR}(D, N)$  is defined in (14) and

$$\tilde{R}_{HB}^{CR}(D_1, D_2) \triangleq p_1(1 - H(D_1)) + p_2(H(D_1) - H(D_2)). \quad (17)$$

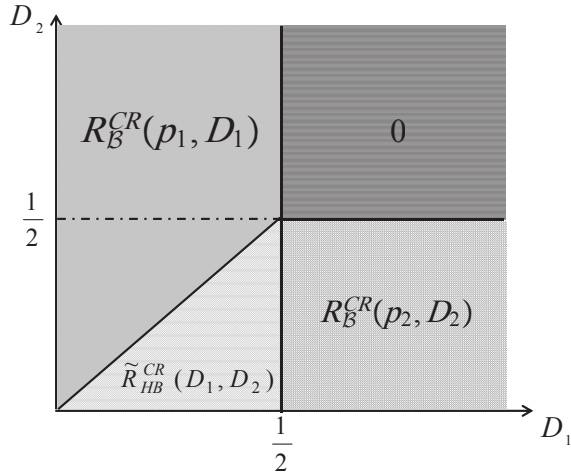


Figure 3. Illustration of the distortion regions in the rate-distortion function (16) for a binary source with degraded erased side information and Hamming distortion.

Moreover, for the same source under erasure distortion the rate-distortion function is given by (16) by substituting  $R_{B\mathcal{E}}^{CR}(D_j, p_j)$  as defined in (15) for  $R_B^{CR}(D_j, p_j)$  for  $j = 1, 2$  and by substituting  $\tilde{R}_{HB,E}^{CR}(D_1, D_2) \triangleq p_1(1 - D_1) + p_2(D_1 - D_2)$  for (17).

Characterization of the rate distortion function (16) requires different considerations for the four subregions of the  $(D_1, D_2)$  plane sketched in Fig. 3. In fact, for  $D_1 \geq 1/2$  and  $D_2 \geq 1/2$ , the required rate is zero, since the distortion constraints are trivially met by setting  $\hat{X}_1 = \hat{X}_2 = 0$  in the achievable rate (9). For the case  $D_1 \geq 1/2$  and  $D_2 \leq 1/2$ , it is sufficient to cater only to Decoder 2 by setting  $\hat{X}_1 = 0$  and  $X = \hat{X}_2 \oplus Q_2$ , with  $Q_2 \sim \text{Ber}(D_2)$  independent of  $X$ , in the achievable rate (9). That this rate cannot be improved upon is a consequence of the trivial converse

$$R_{HB}^{CR}(D_1, D_2) \geq \max\{R_B^{CR}(D_1, p_1), R_B^{CR}(D_2, p_2)\}, \quad (18)$$

which follows by cut-set arguments. The same converse suffices also for the regime  $D_1 \leq 1/2$  and  $D_2 \geq \min(D_1, 1/2)$ . For this case, achievability follows by setting  $X = \hat{X}_2 \oplus Q_2$  and  $\hat{X}_1 = \hat{X}_2$  in (9), where  $Q_2 \sim \text{Ber}(D_2)$  is independent of  $X$ . In the remaining case, namely  $D_2 \leq D_1 \leq 1/2$ , the rate-distortion function does not follow from the point-to-point result (14) as for the regimes discussed thus far. The analysis of this case can be found in [4, Appendix D]. Similar arguments apply also for the erasure distortion metric.

We now compare the rate-distortion function above with the following related settings: (i) the Kaspi model [8], in which the encoder knows the side information, and for which the rate-distortion function  $R_{Kaspi}(D_1, D_2)$  for the example at hand was calculated in [7]; and (ii) the HB setting with no CR constraint, whose rate-distortion function  $R_{HB}(D_1, D_2)$  for the example at hand was derived in [6]. We clearly have the inequalities

$$R_{Kaspi}(D_1, D_2) \leq R_{HB}(D_1, D_2) \leq R_{HB}^{CR}(D_1, D_2), \quad (19)$$

where the first inequality in (19) accounts for the impact of the availability of the side information at the encoder, while the second reflects the potential performance loss due to the CR constraint.

Fig. 4 shows the aforementioned rate-distortion functions with  $p_1 = 1$  and  $p_2 = 0.35$ , which corresponds to the case where Decoder 1 has no side information, for two values of the distortion  $D_2$  versus the distortion  $D_1$ . For  $D_2 \geq p/2 = 0.175$ , the given settings reduce to one in which the encoder needs to communicate information only to Decoder 1. Since Decoder 1 has no side information, the Kaspi and HB settings yield the same performance i.e.,  $R_{Kaspi}(D_1, D_2) = R_{HB}(D_1, D_2)$ . Moreover, if  $D_1$  is sufficiently smaller than  $D_2$ , the operation of the encoder is limited by the distortion requirements of Decoder 1. In this case, Decoder 2 can in fact reconstruct as  $\hat{X}_1 = \hat{X}_2$  while still satisfying its distortion constraints. Therefore, we obtain the same performance in all of the three settings, i.e.,  $R_{Kaspi}(D_1, D_2) = R_{HB}(D_1, D_2) = R_{HB}^{CR}(D_1, D_2)$ . We also note the general performance loss due to the CR constraint, unless, as discussed above, distortion  $D_1$  is sufficiently smaller than  $D_2$ .

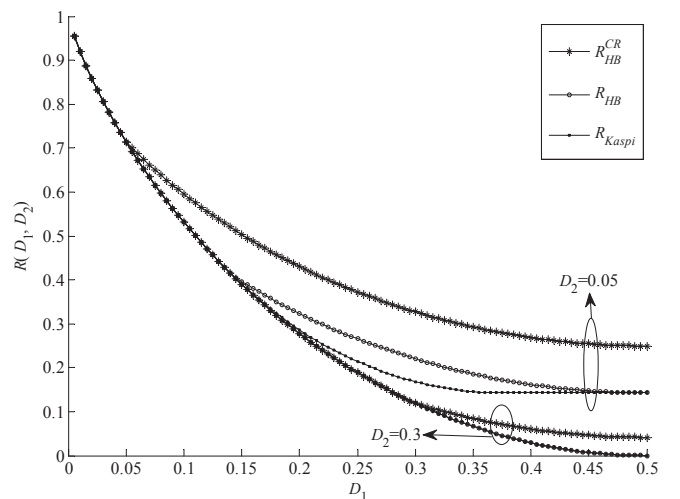


Figure 4. Rate-distortion functions  $R_{Kaspi}(D_1, D_2)$  [7],  $R_{HB}(D_1, D_2)$  [6] and  $R_{HB}^{CR}(D_1, D_2)$  (16) for a binary source under erased side information versus distortion  $D_1$  ( $p_1 = 1$ ,  $p_2 = 0.35$ ,  $D_2 = 0.05$  and  $D_2 = 0.3$ ).

### III. HEGGARD-BERGER PROBLEM WITH COOPERATIVE DECODERS AND COMMON RECONSTRUCTION

The system model for the HB problem with CR and decoder cooperation is similar to the one provided in Sec. II-A with the following differences. Here, in addition to encoding function given in (2) which maps the source sequence  $X^n$  into a message  $J_1$  of  $nR_1$  bits, there is an encoder at Decoder 1 given by  $g_1: [1, 2^{nR_1}] \times \mathcal{Y}_1^n \rightarrow [1, 2^{nR_2}]$ , which maps message  $J_1$  and the source sequence  $Y_1^n$  into a message  $J_2$ . Moreover, instead of the decoding function given in (3), we have the decoding function for Decoder 2,  $h_2: [1, 2^{nR_1}] \times [1, 2^{nR_2}] \times \mathcal{Y}_2^n \rightarrow \hat{\mathcal{X}}_2^n$ , which maps the messages  $J_1$  and  $J_2$  and the side information  $Y_2^n$  into the estimated sequence  $\hat{X}_2^n$ .



### A. Rate-Distortion Region for $X - Y_1 - Y_2$

In this section, a single-letter characterization of the rate-distortion region is derived under the assumption that the joint pmf  $p(x, y_1, y_2)$  is such that the Markov chain relationship  $X - Y_1 - Y_2$  holds.

**Proposition 6.** *The rate-distortion region  $\mathcal{R}^{CR}(D_1, D_2)$  for the HB source coding problem with CR and cooperative decoders under the assumption that  $X - Y_1 - Y_2$  forms a Markov chain is given by the union of all rate pairs  $(R_1, R_2)$  that satisfy the conditions*

$$R_1 \geq I(X; \hat{X}_1 \hat{X}_2 | Y_1) \quad (20a)$$

$$\text{and } R_1 + R_2 \geq I(X; \hat{X}_2 | Y_2) + I(X; \hat{X}_1 | Y_1, \hat{X}_2), \quad (20b)$$

where the mutual information terms are evaluated with respect to the joint pmf

$$p(x, y_1, y_2, \hat{x}_1, \hat{x}_2) = p(x, y_1)p(y_2|y_1)p(\hat{x}_1, \hat{x}_2|x), \quad (21)$$

for some pmf  $p(\hat{x}_1, \hat{x}_2|x)$  such that the constraints (11) are satisfied.

The proof of the converse can be found in [4]. As for the achievability, it follows as a straightforward extension of [3, Sec. III] to the setup at hand where Decoder 2 has side information as well. It is worth emphasizing that the reconstruction  $\hat{X}_2$  for Decoder 2, which has degraded side information, is conveyed by using both the direct link from the Encoder, of rate  $R_1$ , and the path Encoder-Decoder 1-Decoder 2. The latter path leverages the better side information at Decoder 1 and the cooperative link of rate  $R_2$ .

*Remark 7.* If we remove the CR constraint, the problem of determining the rate-distortion region for the setting of Fig. 1 under the Markov assumption  $X - Y_1 - Y_2$  is still open [3].

### B. Rate-Distortion Region for $X - Y_2 - Y_1$

In this section, a single-letter characterization of the rate-distortion region is derived under the assumption that the joint pmf  $p(x, y_1, y_2)$  is such that the Markov chain relationship  $X - Y_2 - Y_1$  holds.

**Proposition 8.** *The rate-distortion region  $\mathcal{R}^{CR}(D_1, D_2)$  for the HB source coding problem with CR and cooperative decoders under the assumption the Markov chain relationship  $X - Y_2 - Y_1$  is given by the union of all rate pairs  $(R_1, R_2)$  that satisfy the conditions*

$$R_1 \geq I(X; \hat{X}_1 | Y_1) + I(X; \hat{X}_2 | Y_2, \hat{X}_1) \quad (22a)$$

$$\text{and } R_2 \geq 0, \quad (22b)$$

where the mutual information terms are evaluated with respect to the joint pmf

$$p(x, y_1, y_2, \hat{x}_1, \hat{x}_2) = p(x, y_2)p(y_1|y_2)p(\hat{x}_1, \hat{x}_2|x), \quad (23)$$

for some pmf  $p(\hat{x}_1, \hat{x}_2|x)$  such that the constraints (11) are satisfied.

The proof of achievability follows immediately by neglecting the link of rate  $R_2$  and using rate  $R_1$  as per the HB

scheme of Proposition 2. The converse follows by considering an enhanced system in which Decoder 2 is provided with the side information of Decoder 1. In this system, link  $R_2$  becomes useless since Decoder 2 possesses all the information available at Decoder 1. It follows that the system reduces to the HB problem with degraded sources studied in the previous section and the bound (22a) follows immediately from Proposition 2.

*Remark 9.* In the case without CR, the rate-distortion function is given similarly to (22), but with the HB rate-distortion function (12) in lieu of the rate-distortion function of the HB problem with CR in (22a).

## IV. CONCLUDING REMARKS

The Common Reconstruction requirement [2], substantially modifies the problem of source coding in the presence of side information at the decoders. A general subject of theoretical interest is identifying those models for which the CR requirements enables a solution of problems that have otherwise resisted solutions for some time. For instance, reference [4] extends the result of this paper to cascade models in the sense of [9]. Other examples that are worth investigating include the Heegard-Berger and cascade source coding problems with no assumptions on side information degradedness and the one-helper lossy source coding problem. Applications of the generalized CR constraint of [10] are also of interest (see [4]).

## ACKNOWLEDGMENT

The work of O. Simeone was supported by the U.S. National Science Foundation under grant CCF-0914899, and the work of H. V. Poor and R. Tandon was supported in part by the U.S. Air Force Office of Scientific Research under MURI Grant FA9550-09-1-0643 and in part by the U.S. National Science Foundation under Grant CNS-09-05398.

## REFERENCES

- [1] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Trans. Inform. Theory*, vol. 31, no. 6, pp.727–734, Nov. 1985.
- [2] Y. Steinberg, "Coding and common reconstruction," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, 2009.
- [3] D. Vasudevan, "On the Heegard-Berger/Kaspi problem with decoder cooperation," in *Proc. IEEE Int'l Conf. on Telecommunications*, St. Petersburg, Russia, June 2008.
- [4] B. Ahmadi, R. Tandon, O. Simeone and H. V. Poor, "Heegard-Berger and cascade source coding problems with common reconstruction constraints," submitted [arXiv:1112.1762].
- [5] T. Weissman and A. El Gamal, "Source coding with limited-look-ahead side information at the decoder," *IEEE Trans. on Inform. Theory*, vol. 52, no. 12, pp. 5218–5239, Dec. 2006.
- [6] R. Tandon, S. Mohajer, and H. V. Poor, "Cascade source coding with erased side information," in *Proc. IEEE Int'l Symp. Inform. Theory*, St. Petersburg, Russia, Aug. 2011.
- [7] E. Perron, S. N. Diggavi, and E. Telatar, "Lossy source coding with Gaussian or erased side-information," in *Proc. IEEE Int'l Symp. Inform. Theory*, Seoul, South Korea, Jul. 2009.
- [8] A. H. Kaspi, "Rate-distortion function when side-information may be present at the decoder," *IEEE Trans. Inform. Theory*, vol. 40, pp. 2031–2034, Nov. 1994.
- [9] Y. K. Chia, H. Permuter and T. Weissman, "Cascade, triangular and two way source coding with degraded side information at the second user," submitted [http://arxiv.org/abs/1010.3726].
- [10] A. Lapidoto, A. Malar, and M. Wigger, "Constrained Wyner-Ziv coding," in *Proc. IEEE Int'l Symp. Inform. Theory*, St. Petersburg, Russia, Aug. 2011.