

On Lossy Transmission of Correlated Sources over a Multiple Access Channel

Başak Güler

The Pennsylvania State University
University Park, PA
basak@psu.edu

Deniz Gündüz

Imperial College London
London, UK
d.gunduz@imperial.ac.uk

Aylin Yener

The Pennsylvania State University
University Park, PA
yener@ee.psu.edu

Abstract—We study lossy communication of correlated sources over a multiple access channel. In particular, we provide a joint source-channel coding scheme for transmitting correlated sources with decoder side information, and study the conditions under which separate source and channel coding is optimal. For the latter, the encoders and/or the decoder have access to a common observation conditioned on which the two sources are independent. By establishing necessary and sufficient conditions, we show the optimality of separation when the encoders and the decoder both have access to the common observation. We also demonstrate that separation is optimal when only the encoders have access to the common observation whose lossless recovery is required at the decoder. As a special case, we study separation for sources with a common part. Our results indicate that side information can have significant impact on the optimality of source-channel separation in lossy transmission.

I. INTRODUCTION

We consider the transmission of two correlated memoryless sources over a multiple access channel with fidelity criteria. The encoders and/or the decoder may have access to side information correlated with the sources. We propose an achievable joint source-channel coding scheme in the presence of correlated decoder side information. We then focus on the case when the two sources are conditionally independent given the side information available at the encoders and/or the decoder. First, we identify the necessary and sufficient conditions under which separation is optimal when the side information is shared between the encoders and the decoder. Next, we consider the case when the decoder is required to recover some common observation shared by both encoders losslessly, but can tolerate some distortion for the parts known only at a single encoder. We show that separation is also optimal for this case. Lastly, we consider the transmission of sources with a common part in the sense of Gács-Körner [1], and investigate the conditions under which separation is optimal in the absence of side information.

Related Work: Shannon proved the optimality of separate source and channel coding for transmitting a source through a noisy channel [2], known as the separation theorem. Separation was shown to be optimal for the lossy transmission of a source with decoder side information, in [3]. The point-to-point scenario was extended in [4] to transmission of correlated sources through a multiple access channel, and separation

This research is in part sponsored by the U.S. Army Research Laboratory under the Network Science Collaborative Technology Alliance, Agreement Number W911NF-09-2-0053.

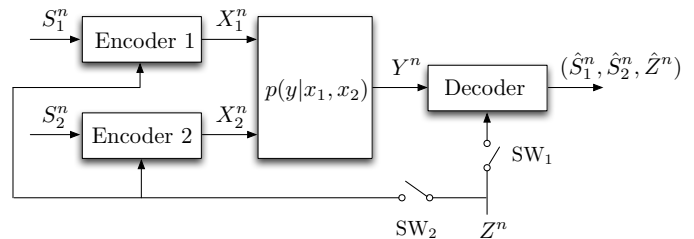


Fig. 1. Communication of correlated sources over a multiple access channel.

was shown to be optimal when one of the sources is shared between the two encoders. For the lossless case, the optimality of separation was established in [5] for transmitting correlated sources through a multiple access channel, whenever the decoder has access to some side information conditioned on which the two sources are independent. A joint source-channel coding scheme was proposed in [6] for the transmission of correlated sources over a multiple access channel based on hybrid coding.

In the remainder of the paper, X is a random variable, and x is its realization. $X^n = (X_1, \dots, X_n)$ is a random vector of length n , and $x^n = (x_1, \dots, x_n)$ denotes its realization. \mathcal{X} is a set with cardinality $|\mathcal{X}|$. $\mathbb{E}[X]$ is the expected value of X .

II. SYSTEM MODEL

Consider the transmission of two discrete memoryless sources S_1 and S_2 in Fig. 1. Encoder 1 observes $S_1^n = (S_{11}, \dots, S_{1n})$. Encoder 2 observes $S_2^n = (S_{21}, \dots, S_{2n})$. If switch SW_2 in Fig. 1 is closed, the two encoders also have access to a common observation Z^n correlated with S_1^n and S_2^n . Encoders 1 and 2 map their observations to the channel inputs X_1^n and X_2^n , respectively. A discrete memoryless multiple access channel (DM-MAC) exists between the encoders and the decoder, characterized by the distribution $p(y|x_1, x_2)$. If switch SW_1 in Fig. 1 is closed, the decoder has access to side information Z^n . Upon observing the channel output Y^n and side information Z^n whenever it is available, the decoder constructs \hat{S}_1^n , \hat{S}_2^n , and \hat{Z}^n such that

$$\frac{1}{n} \sum_{i=1}^n \mathbb{E}[d_j(S_{ji}, \hat{S}_{ji})] \leq D_j \text{ for } j = 1, 2 \quad (1)$$

where D_j is the maximum average distortion allowed for S_j , given a distortion measure $d_j(s_{ji}, \hat{s}_{ji})$ for $j = 1, 2$, and $P(Z^n \neq \hat{Z}^n) \rightarrow 0$ as $n \rightarrow \infty$. Random variables $S_1, S_2, Z, X_1, X_2, Y, \hat{S}_1, \hat{S}_2, \hat{Z}$ are defined over the corresponding

alphabets $\mathcal{S}_1, \mathcal{S}_2, \mathcal{Z}, \mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}, \hat{\mathcal{S}}_1, \hat{\mathcal{S}}_2, \hat{\mathcal{Z}}$. Note that, when the switch SW_1 is closed, error probability in decoding Z^n becomes irrelevant since it is readily available at the decoder, and serves as side information.

We use the following notation from [3], [4]. Define the minimum average distortion for S_j given Q as

$$\mathcal{E}(S_j|Q) = \min_{f:Q \rightarrow \hat{\mathcal{S}}_j} E[d_j(S_j, f(Q))], \quad j = 1, 2, \quad (2)$$

and the conditional rate distortion function [7] for source S_j when side information Z is shared between the encoder and the decoder as

$$R_{S_j|Z}(D_j) = \min_{\substack{p(u_j|s_j, z) \\ \mathcal{E}(S_j|U_j, Z) \leq D_j}} I(S_j; U_j|Z), \quad j = 1, 2. \quad (3)$$

III. JOINT SOURCE-CHANNEL CODING WITH DECODER SIDE INFORMATION

We first assume that only SW_1 is closed, and present a general achievable scheme for the lossy communication of correlated sources in the presence of decoder side information.

Theorem 1. *The distortion pair (D_1, D_2) is achievable for sending two discrete memoryless correlated sources S_1 and S_2 over a DM-MAC with $p(y|x_1, x_2)$ and decoder side information Z if there exists a joint distribution $p(u_1, u_2, s_1, s_2, z) = p(u_1|s_1)p(u_2|s_2)p(s_1, s_2, z)$, and functions $x_j(u_j, s_j)$, $g_j(u_1, u_2, y, z)$ for $j = 1, 2$, such that*

$$I(U_1; S_1|U_2, Z) < I(U_1; Y|U_2, Z) \quad (4)$$

$$I(U_2; S_2|U_1, Z) < I(U_2; Y|U_1, Z) \quad (5)$$

$$I(U_1, U_2; S_1, S_2|Z) < I(U_1, U_2; Y|Z) \quad (6)$$

where $\mathbb{E}[d_j(S_j, g_j(U_1, U_2, Y, Z))] \leq D_j$ for $j = 1, 2$.

Proof. Our achievable scheme builds upon the hybrid coding framework of [6], by generalizing it to the case with decoder side information. \square

For the remainder of the paper, we assume that the sources are independent when conditioned on the side information, i.e., the Markov condition $S_1 - Z - S_2$ holds.

A. Optimality of Separation

We now focus on the conditions under which separation is optimal when sources S_1 and S_2 are independent given the side information Z . We first consider that both switches in Fig. 1 are closed. We show that whenever the two sources are independent given the side information that is shared between the encoders and the decoder, separation is optimal. The next theorem states the necessary and sufficient conditions.

Theorem 2. *Consider the communication of two correlated sources S_1 and S_2 with side information Z shared between the encoders and the decoder. If $S_1 - Z - S_2$, then separate source and channel coding is optimal, and a distortion pair (D_1, D_2) is achievable if*

$$R_{S_1|Z}(D_1) < I(X_1; Y|X_2, Q) \quad (7)$$

$$R_{S_2|Z}(D_2) < I(X_2; Y|X_1, Q) \quad (8)$$

$$R_{S_1|Z}(D_1) + R_{S_2|Z}(D_2) < I(X_1, X_2; Y|Q) \quad (9)$$

for some $p(x_1, x_2, y, q) = p(y|x_1, x_2)p(x_1|q)p(x_2|q)p(q)$.

Conversely, for any achievable (D_1, D_2) pair, (7)-(9) should hold with $<$ replaced with \leq .

Proof. (Achievability) The source coding part is based on lossy source coding at the two encoders conditioned on side information Z shared between the encoder and decoder [7], after which the conditional rate distortion functions given in (3) can be achieved for S_1 and S_2 , respectively. Channel coding is performed based on classical multiple access channel coding with independent channel inputs [8].

(Converse) Suppose there exist encoding functions $e_j : \mathcal{S}_j^n \times \mathcal{Z}^n \rightarrow \mathcal{X}_j^n$ for encoder $j = 1, 2$, and decoding functions $g_j : \mathcal{Y}^n \times \mathcal{Z}^n \rightarrow \hat{\mathcal{S}}_j^n$ such that $\frac{1}{n} \sum_{i=1}^n E[d_j(S_{ji}, \hat{S}_{ji})] \leq D_j + \epsilon$ for $j = 1, 2$, where $\epsilon \rightarrow 0$ as $n \rightarrow \infty$.

Define $U_{ji} = (Y^n, S_j^{i-1}, Z_i^c)$ for $j = 1, 2$, where $Z_i^c = (Z_1, \dots, Z_{i-1}, Z_{i+1}, \dots, Z_n)$. Then,

$$\frac{1}{n} I(X_1^n; Y^n | X_2^n, Z^n) \quad (10)$$

$$\geq \frac{1}{n} I(S_1^n; Y^n | X_2^n, Z^n) \quad (11)$$

$$= \frac{1}{n} I(S_1^n; Y^n, X_2^n | Z^n) \quad (12)$$

$$\geq \frac{1}{n} I(S_1^n; Y^n | Z^n) \quad (13)$$

$$= \frac{1}{n} \sum_{i=1}^n I(S_{1i}; Y^n | S_1^{i-1}, Z^n) \quad (14)$$

$$= \frac{1}{n} \sum_{i=1}^n (I(S_{1i}; Y^n, S_1^{i-1}, Z_i^c | Z_i) - I(S_{1i}; S_1^{i-1}, Z_i^c | Z_i)) \quad (15)$$

$$= \frac{1}{n} \sum_{i=1}^n I(S_{1i}; U_{1i} | Z_i) \quad (16)$$

$$= \frac{1}{n} \sum_{i=1}^n R_{S_1|Z}(\mathcal{E}(S_{1i} | U_{1i}, Z_i)) \quad (17)$$

$$\geq \frac{1}{n} \sum_{i=1}^n R_{S_1|Z}(\mathcal{E}(S_{1i} | Z^n, Y^n)) \quad (18)$$

$$\geq \frac{1}{n} \sum_{i=1}^n R_{S_1|Z}(E[d_1(S_{1i}, \hat{S}_{1i})]) \quad (19)$$

$$\geq R_{S_1|Z}(D_1 + \epsilon) \quad (20)$$

(11) is from $Y^n - X_1^n X_2^n - S_1^n Z^n$ and conditioning cannot increase entropy, and (12) from $X_2^n - Z^n - S_1^n$:

$$p(x_2^n, s_1^n | z^n) = \sum_{s_2^n} p(x_2^n, s_1^n, s_2^n | z^n) \quad (21)$$

$$= \sum_{s_2^n} p(x_2^n | s_1^n, s_2^n, z^n) p(s_1^n | z^n) p(s_2^n | z^n) \quad (22)$$

$$= p(s_1^n | z^n) \sum_{s_2^n} p(x_2^n | s_2^n, z^n) p(s_2^n | z^n) \quad (23)$$

$$= p(x_2^n | z^n) p(s_1^n | z^n) \quad (24)$$

where (22) is from $S_1^n - Z^n - S_2^n$; (23) is from $X_2^n - S_2^n Z^n - S_1^n$; (13) is from the nonnegativity of mutual information; (14) is from the chain rule; (16) is from the memoryless property of

the sources and the side information; (17) is from (2) and (3); (18) is from the fact that further conditioning cannot increase (2); (19) follows since \hat{S}_{1i} is a function of (Y^n, Z^n) ; (20) holds as $R_{S_1|Z}(D_1)$ is convex and monotone in D_1 .

By defining a discrete random variable \tilde{Q} uniformly distributed over $\{1, \dots, n\}$ independent of everything else,

$$\frac{1}{n}I(X_1^n; Y^n | X_2^n, Z^n) \quad (25)$$

$$\leq \frac{1}{n} \sum_{i=1}^n (H(Y_i | X_{2i}, Z^n) - H(Y_i | X_{1i}, X_{2i}, Z^n)) \quad (26)$$

$$= \frac{1}{n} \sum_{i=1}^n I(X_{1i}; Y_i | X_{2i}, \tilde{Q} = i, Z^n) \quad (27)$$

$$= I(X_{1\tilde{Q}}; Y_{\tilde{Q}} | X_{2\tilde{Q}}, \tilde{Q}, Z^n) = I(X_1; Y | X_2, Q) \quad (28)$$

where we let $X_1 = X_{1\tilde{Q}}$, $X_2 = X_{2\tilde{Q}}$, $Y = Y_{\tilde{Q}}$ and $Q = (\tilde{Q}, Z^n)$. Combining (10) and (20) with (28) yields (7).

Following similar steps, we obtain (8). Lastly, we have

$$\begin{aligned} \frac{1}{n}I(X_1^n, X_2^n; Y^n | Z^n) &= \frac{1}{n}(I(X_1^n; Y^n | X_2^n, Z^n) + I(X_2^n; Y^n | Z^n)) \\ &\geq R_{S_1|Z}(D_1 + \epsilon) + \frac{1}{n}I(S_2^n; Y^n | Z^n) \quad (29) \\ &\geq R_{S_1|Z}(D_1 + \epsilon) + R_{S_2|Z}(D_2 + \epsilon) \quad (30) \end{aligned}$$

the first term in (29) is from (10)-(20), and (30) is from (13)-(20) with the role of S_1 replaced with S_2 . To obtain the second term in (29), we first show that $Y^n - Z^n X_2^n - S_2^n$:

$$\begin{aligned} p(y^n, s_2^n | z^n, x_2^n) \\ = p(s_2^n | z^n, x_2^n) p(y^n | s_2^n, z^n, x_2^n) \quad (31) \end{aligned}$$

$$= p(s_2^n | z^n, x_2^n) \sum_{s_1^n, x_1^n} p(y^n, s_1^n, x_1^n | s_2^n, z^n, x_2^n) \quad (32)$$

$$= p(s_2^n | z^n, x_2^n) \sum_{s_1^n, x_1^n} p(y^n | x_1^n, x_2^n) p(x_1^n | s_1^n, z^n) p(s_1^n | z^n) \quad (33)$$

$$= p(s_2^n | z^n, x_2^n) \sum_{x_1^n} p(y^n | x_1^n, x_2^n) p(x_1^n | z^n) \quad (34)$$

(33) is from $Y^n - X_1^n X_2^n - S_1^n S_2^n Z^n$ and $X_1^n - S_1^n Z^n - S_2^n X_2^n$ as well as $S_1^n - Z^n - S_2^n X_2^n$ which holds since

$$p(s_1^n, s_2^n, x_2^n | z^n) = p(x_2^n | s_1^n, s_2^n, z^n) p(s_2^n | z^n) p(s_1^n | z^n) \quad (35)$$

$$= p(x_2^n | s_2^n, z^n) p(s_2^n | z^n) p(s_1^n | z^n) \quad (36)$$

$$= p(x_2^n, s_2^n | z^n) p(s_1^n | z^n) \quad (37)$$

(35) is from $S_1^n - Z^n - S_2^n$ and (36) is from $X_2^n - S_2^n Z^n - S_1^n$. Note that

$$p(y^n | z^n, x_2^n) = \sum_{s_1^n, x_1^n} p(y^n, s_1^n, x_1^n | z^n, x_2^n) \quad (38)$$

$$= \sum_{s_1^n, x_1^n} p(y^n | x_1^n, x_2^n) p(x_1^n | s_1^n, z^n) p(s_1^n | z^n) \quad (39)$$

$$= \sum_{x_1^n} p(y^n | x_1^n, x_2^n) p(x_1^n | z^n). \quad (40)$$

(39) is from $Y^n - X_1^n X_2^n - S_1^n S_2^n Z^n$ and $X_1^n - S_1^n Z^n - X_2^n$ as well as $S_1^n - Z^n - X_2^n$ which holds since $S_1^n - Z^n - S_2^n X_2^n$. Combining (40) and (34), we have

$$p(y^n, s_2^n | z^n, x_2^n) = p(s_2^n | z^n, x_2^n) p(y^n | z^n, x_2^n) \quad (41)$$

hence $Y^n - Z^n X_2^n - S_2^n$. Then, (29) follows from,

$$I(X_2^n; Y^n | Z^n) = H(Y^n | Z^n) - H(Y^n | X_2^n, Z^n) \quad (42)$$

$$= H(Y^n | Z^n) - H(Y^n | X_2^n, Z^n, S_2^n) \quad (43)$$

$$\geq H(Y^n | Z^n) - H(Y^n | Z^n, S_2^n) \quad (44)$$

$$= I(S_2^n; Y^n | Z^n) \quad (45)$$

(43) is from $Y^n - Z^n X_2^n - S_2^n$ and (44) holds as conditioning cannot increase entropy. Lastly,

$$\begin{aligned} \frac{1}{n}I(X_1^n, X_2^n; Y^n | Z^n) \\ \leq \frac{1}{n} \sum_{i=1}^n (H(Y_i | Z^n) - H(Y_i | X_{1i}, X_{2i}, Z^n)) \quad (46) \end{aligned}$$

$$= \frac{1}{n} \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_i | \tilde{Q} = i, Z^n) = I(X_1, X_2; Y | Q) \quad (47)$$

Combining (30) with (47) recovers (9).

We show $p(x_1, x_2 | q) = p(x_1 | q) p(x_2 | q)$ for $q = (i, z^n)$,

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2 | Q = q) \\ = P(X_{1i} = x_1, X_{2i} = x_2 | \tilde{Q} = i, Z^n = z^n) \quad (48) \end{aligned}$$

$$= P(X_{1i} = x_1 | \tilde{Q} = i, Z^n = z^n) P(X_{2i} = x_2 | \tilde{Q} = i, Z^n = z^n) \quad (49)$$

$$= P(X_1 = x_1 | Q = q) P(X_2 = x_2 | Q = q) \quad (50)$$

as in [5], (49) is from $X_{1i} - Z^n - X_{2i}$ as $S_1^n - Z^n - S_2^n$. \square

For the case when side information Z is available only at the decoder, i.e., when only switch SW_1 is closed, separation is known to be optimal for the lossless transmission of sources S_1 and S_2 whenever $S_1 - Z - S_2$ [5]. In light of Theorem 2, we show that a similar result holds for the lossy case whenever the Wyner-Ziv rate distortion function of each source is equal to its conditional rate distortion function.

Corollary 1. Consider communication of two correlated sources S_1 and S_2 with decoder only side information Z . If

$$R_{S_j|Z}(D_j) = R_{S_j|Z}^{\text{WZ}}(D_j) \quad (51)$$

and $S_1 - Z - S_2$, where

$$R_{S_j|Z}^{\text{WZ}}(D_j) = \min_{\substack{p(u_j | s_j), g(u_j, z) \\ \mathbb{E}[d_j(S_j, g(U_j, Z))] \leq D_j \\ U_j - S_j - Z}} I(S_j; U_j | Z) \text{ for } j = 1, 2$$

is the (Wyner-Ziv) rate distortion function of S_j with decoder side information Z [9], then separation is optimal, with the necessary and sufficient conditions stated as in (7)-(9).

Corollary 1 follows from the fact that whenever (51) holds, conditional rate distortion functions in Theorem 2 are achievable by relying only on decoder side information.

Remark 1. Gaussian sources are an example for (51).

IV. SEPARATION IN THE PRESENCE OF COMMON OBSERVATION

In this section, we assume that only switch SW_2 is closed in Fig. 1, and show the optimality of separation whenever lossless reconstruction of the common observation is required.

Theorem 3. Consider the communication of correlated sources S_1, S_2 , and Z where Z is observed by both encoders. If $S_1 - Z - S_2$ and that Z is to be reconstructed in a lossless fashion, then, separate source and channel coding is optimal, and the distortion pair (D_1, D_2) is achievable if

$$R_{S_1|Z}(D_1) < I(X_1; Y|X_2, W) \quad (52)$$

$$R_{S_2|Z}(D_2) < I(X_2; Y|X_1, W) \quad (53)$$

$$R_{S_1|Z}(D_1) + R_{S_2|Z}(D_2) < I(X_1, X_2; Y|W) \quad (54)$$

$$H(Z) + R_{S_1|Z}(D_1) + R_{S_2|Z}(D_2) < I(X_1, X_2; Y) \quad (55)$$

for some $p(x_1, x_2, y, w) = p(y|x_1, x_2)p(x_1|w)p(x_2|w)p(w)$.

Conversely, if a distortion pair (D_1, D_2) is achievable, then (52)-(55) should hold with $<$ replaced with \leq .

Proof. We provide an outline of our proof in the sequel.

(Achievability) The source coding part is based on [10, Theorem 1]. We let $Y_0 \leftarrow Z$, $Y_j \leftarrow (S_j, Z)$ for $j = 1, 2$, and $X \leftarrow Z$ in [10, Fig. 3] and observe that any achievable rate for this system is also achievable by our system, since Y_0 is known by both encoders in our system as $Y_0 \leftarrow Z$. Hence, the encoders can cooperate to transmit Y_0 and realize any achievable scheme in [10].

Letting $U = X$ in [10, Theorem 1] and substituting $X \leftarrow Z$, $Y_0 \leftarrow Z$, $\hat{Y}_0 \leftarrow \hat{Z}$, $Y_j \leftarrow (S_j, Z)$, $V_j \leftarrow U_j$, $\hat{Y}_j \leftarrow \hat{S}_j$, and $d_j(Y_j, \hat{Y}_j) \leftarrow d_j(S_j, \hat{S}_j)$ for $j = 1, 2$, we can show that a distortion pair (D_1, D_2) is achievable for the rate triplet (R_0, R_1, R_2) if

$$R_1 \geq I(S_1; U_1|Z) \quad (56)$$

$$R_2 \geq I(S_2; U_2|Z) \quad (57)$$

$$R_1 + R_2 \geq I(S_1; U_1|Z) + I(S_2; U_2|Z) \quad (58)$$

$$R_0 + R_1 + R_2 \geq H(Z) + I(S_1; U_1|Z) + I(S_2; U_2|Z) \quad (59)$$

and $\mathbb{E}[d_j(S_j, \hat{S}_j)] \leq D_j$ for $j = 1, 2$, for some distribution

$$p(z, s_1, s_2)p(u_1|s_1, z)p(u_2|s_2, z)p(\hat{s}_1, \hat{s}_2|z, u_1, u_2). \quad (60)$$

We then show that one can set $\hat{S}_j = f_j(Z, U_1, U_2)$ for $j = 1, 2$ without loss of optimality, and that for any function $f_j(Z, U_1, U_2)$ such that $\mathbb{E}[d_j(S_j, f_j(Z, U_1, U_2))] \leq D_j$, there is a function $g_j(Z, U_j)$ with $\mathbb{E}[d_j(S_j, g_j(Z, U_j))] \leq \mathbb{E}[d_j(S_j, f_j(Z, U_1, U_2))] \leq D_j$, along the lines of [11]. Using this result, we restate (56)-(59) as follows.

A distortion pair (D_1, D_2) is achievable for (R_0, R_1, R_2) if

$$R_1 \geq R_{S_1|Z}(D_1) \quad (61)$$

$$R_2 \geq R_{S_2|Z}(D_2) \quad (62)$$

$$R_1 + R_2 \geq R_{S_1|Z}(D_1) + R_{S_2|Z}(D_2) \quad (63)$$

$$R_0 + R_1 + R_2 \geq H(Z) + R_{S_1|Z}(D_1) + R_{S_2|Z}(D_2) \quad (64)$$

since for any $p(s_j, u_j, z) = p(u_j|s_j, z)p(s_j|z)p(z)$ and $g_j(z, u_j)$ such that $\mathbb{E}[d_j(S_j, g_j(Z, U_j))] \leq D_j$,

$$I(S_j; U_j|Z) \geq R_{S_j|Z}(D_j), \quad j = 1, 2, \quad (65)$$

with $R_{S_j|Z}(D_j)$ from (3), completing the source coding part.

Our channel coding scheme is based on multiple access channel coding with a common message [12], for which any triplet of rates (R_0, R_1, R_2) is achievable if

$$R_1 \leq I(X_1; Y|X_2, W) \quad (66)$$

$$R_2 \leq I(X_2; Y|X_1, W) \quad (67)$$

$$R_1 + R_2 \leq I(X_1, X_2; Y|W) \quad (68)$$

$$R_0 + R_1 + R_2 \leq I(X_1, X_2; Y) \quad (69)$$

for a $p(x_1, x_2, y, w) = p(y|x_1, x_2)p(x_1|w)p(x_2|w)p(w)$.

(Converse) Our proof is along the lines of [3] and [4]. Suppose there exist encoding functions $e_j : \mathcal{S}_j^n \times \mathcal{Z}^n \rightarrow \mathcal{X}_j^n$ for $j = 1, 2$, decoding functions $g_j : \mathcal{Y}^n \rightarrow \hat{\mathcal{S}}_j^n$ for $j = 1, 2$ and $g_0 : \mathcal{Y}^n \rightarrow \hat{\mathcal{Z}}^n$ such that $\frac{1}{n} \sum_{i=1}^n \mathbb{E}[d_j(S_{ji}, \hat{S}_{ji})] \leq D_j + \epsilon$ for $j = 1, 2$ and $P(Z^n \neq \hat{Z}^n) \leq P_e$ where $\epsilon \rightarrow 0$, $P_e \rightarrow 0$ as $n \rightarrow \infty$. Define $U_{ji} = (Y^n, S_j^{i-1}, Z_i^c)$ for $j = 1, 2$ where $Z_i^c = (Z_1, \dots, Z_{(i-1)}, Z_{(i+1)}, \dots, Z_n)$. Then,

$$\frac{1}{n} I(X_1^n; Y^n | X_2^n, Z^n) \quad (70)$$

$$\geq \frac{1}{n} I(S_1^n; Y^n | X_2^n, Z^n) \quad (71)$$

$$= \frac{1}{n} I(S_1^n; Y^n, X_2^n | Z^n) \quad (72)$$

$$\geq \frac{1}{n} I(S_1^n; Y^n | Z^n) \quad (73)$$

$$\geq \frac{1}{n} \sum_{i=1}^n R_{S_{1i}|Z}(\mathcal{E}(S_{1i}|U_{1i}, Z_i)) \quad (74)$$

$$\geq \frac{1}{n} \sum_{i=1}^n R_{S_{1i}|Z}(\mathcal{E}(S_{1i}|Y^n)) \quad (75)$$

$$\geq \frac{1}{n} \sum_{i=1}^n R_{S_{1i}|Z}(\mathbb{E}[d_1(S_{1i}, \hat{S}_{1i})]) \geq R_{S_1|Z}(D_1 + \epsilon) \quad (76)$$

(71) is from $Y^n - X_1^n X_2^n - Z^n S_1^n$ and that conditioning cannot increase entropy, and (72) is from $I(S_1^n; X_2^n | Z^n) = 0$ since $S_1^n - Z^n - X_2^n$ as follows.

$$p(x_2^n, s_1^n | z^n) = \sum_{s_2^n} p(x_2^n, s_2^n, s_1^n | z^n) \quad (77)$$

$$= \sum_{s_2^n} p(x_2^n | s_2^n, z^n) p(s_2^n | z^n) p(s_1^n | z^n) \quad (78)$$

$$= p(x_2^n | z^n) p(s_1^n | z^n) \quad (79)$$

(78) holds since $X_2^n - S_2^n Z^n - S_1^n$ and $S_1^n - Z^n - S_2^n$. Equation (74) is from (2) and (3); (75) is from the fact that conditioning cannot increase (2); (76) follows as \hat{S}_{1i} is a function of Y^n and that $R_{S_{1i}|Z}(D_1)$ is convex and monotone in D_1 .

By replacing Q with $W = (\tilde{Q}, Z^n)$ in (25)-(28) we have

$$\frac{1}{n} I(X_1^n; Y^n | X_2^n, Z^n) \leq I(X_1; Y|X_2, W) \quad (80)$$

which, when combined with (70) and (76), leads to (52).

Similar steps lead to (53). Next, we show that

$$\frac{1}{n} I(X_1^n, X_2^n; Y^n) \quad (81)$$

$$\geq \frac{1}{n} I(S_1^n, S_2^n, Z^n; Y^n) \quad (82)$$

$$= \frac{1}{n} (I(Z^n; Y^n) + I(S_1^n; Y^n | Z^n) + H(S_2^n | Z^n) - H(S_2^n | Y^n, S_1^n, Z^n)) \quad (83)$$

$$\geq \frac{1}{n}(I(Z^n; Y^n) + I(S_1^n; Y^n | Z^n) + I(S_2^n; Y^n | Z^n)) \quad (84)$$

$$\geq H(Z) + R_{S_1|Z}(D_1 + \epsilon) + R_{S_2|Z}(D_2 + \epsilon) - \delta(P_e) \quad (85)$$

(82) is from $Y^n - X_1^n X_2^n - S_1^n S_2^n Z^n$, (83) from $S_2^n - Z^n - S_1^n$, (85) from (73)-(76), the memoryless property of Z^n , and Fano's inequality,

$$H(Z^n | Y^n) \leq H(Z^n | \hat{Z}^n) \leq n\delta(P_e) \quad (86)$$

where $\delta(P_e) \rightarrow 0$ as $P_e \rightarrow 0$ [8]. We lastly show that

$$\frac{1}{n}I(X_1^n, X_2^n; Y^n) \leq \frac{1}{n} \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_i | \tilde{Q} = i) \quad (87)$$

$$\leq I(X_1, X_2; Y | \tilde{Q}) \quad (88)$$

$$\leq H(Y) - H(Y | X_1, X_2) = I(X_1, X_2; Y)$$

which, when combined with (85), leads to (55). From similar steps, we can derive (54), and show that $p(x_1, x_2 | w) = p(x_1 | w)p(x_2 | w)$ as in (50) and complete the proof. \square

Corollary 2. *A special case of Theorem 3 is the transmission of two sources over a DM-MAC with one distortion criterion, when one source is available at both encoders as considered in [4], which corresponds to when S_2 is a constant in Theorem 3.*

V. SEPARATION FOR SOURCES WITH A COMMON PART

In this section, we assume that both switches in Fig. 1 are open, $Z = \text{constant}$, and study the conditions under which separate source and channel coding is optimal for transmitting sources with a common part. We first review two notions that quantify common information between correlated sources.

Definition 1. (*Gács-Körner common information*) [1], [13] Define the function $f_j : \mathcal{S}_j \rightarrow \{1, \dots, k\}$ for $j = 1, 2$ with the largest integer k such that $P(f_j(S_j) = u_0) > 0$ for $u_0 \in \{1, \dots, k\}$, $j = 1, 2$ and $P(f_1(S_1) = f_2(S_2)) = 1$. Then, $U_0 = f_1(S_1) = f_2(S_2)$ is defined as the common part between S_1 and S_2 , and the Gács-Körner common information is

$$C_{GK}(S_1, S_2) = H(U_0). \quad (89)$$

Definition 2. (*Wyner's common information*) [14] Wyner's common information between S_1 and S_2 is defined as,

$$C_W(S_1, S_2) = \min_{\substack{p(v|s_1, s_2) \\ S_1 - V - S_2}} I(S_1, S_2; V). \quad (90)$$

Remark 2. $C_{GK}(S_1, S_2) = C_W(S_1, S_2)$ if and only if there exists a random variable U_0 such that U_0 is the common part of S_1 and S_2 from Definition 1, and $S_1 - U_0 - S_2$, [1], [14].

We first state a separation result when common information is available as side information at the decoder.

Corollary 3. *Consider the transmission of two sources S_1 and S_2 with a common part $U_0 = f_1(S_1) = f_2(S_2)$ from Definition 1. Then, separation is optimal whenever*

$$C_{GK}(S_1, S_2) = C_W(S_1, S_2), \quad (91)$$

and the common part U_0 is available at the decoder.

We know from Remark 2 that whenever (91) holds, then $S_1 - U_0 - S_2$, where U_0 is the common part of S_1 and S_2 as in Definition 1. Since the two encoders can extract

U_0 individually, each encoder can achieve the corresponding conditional rate distortion function in which U_0 is shared between the encoder and the decoder. Corollary 3 then follows from Theorem 2 by letting $Z \leftarrow U_0$.

Our next result states that whenever Gács-Körner common information between two sources is equal to Wyner's common information, then separate source and channel coding is optimal if lossless reconstruction of the common part is required.

Corollary 4. *Consider the transmission of two correlated sources S_1 and S_2 with a common part $U_0 = f_1(S_1) = f_2(S_2)$ from Definition 1. Let $C_{GK}(S_1, S_2) = C_W(S_1, S_2)$ and the common part U_0 of S_1 and S_2 is to be recovered losslessly. Then, separate source and channel coding is optimal.*

We have from Definition 1 that the two encoders can separately reconstruct U_0 , and from Remark 2 that $S_1 - U_0 - S_2$. Then, the result follows from letting $Z \leftarrow U_0$ in Theorem 3.

VI. CONCLUSION

We have considered lossy transmission of correlated sources over a multiple access channel. We have investigated the conditions under which separate source and channel coding is optimal when the encoder and/or decoder has access to side information. Current and future directions include identifying further multiple access scenarios for which separation is (sub)optimal as well as other multi-terminal scenarios with side information.

REFERENCES

- [1] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems of Control and Inf. Theory*, vol. 2, no. 2, pp. 149–162, 1973.
- [2] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 625–656, July, Oct. 1948.
- [3] S. Shamai, S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. on Inf. Theory*, vol. 44, no. 2, pp. 564–579, 1998.
- [4] D. Gündüz and E. Erkip, "Correlated sources over an asymmetric multiple access channel with one distortion criterion," in *41st Annual Conf. on Inf. Sciences and Systems, CISS'07*, 2007, pp. 325–330.
- [5] D. Gündüz, E. Erkip, A. Goldsmith, and H. V. Poor, "Source and channel coding for correlated sources over multiuser channels," *IEEE Trans. on Inf. Theory*, vol. 55, no. 9, pp. 3927–3944, 2009.
- [6] P. Minero, S. H. Lim, and Y.-H. Kim, "A unified approach to hybrid coding," *IEEE Trans. on Inf. Theory*, vol. 61, no. 4, pp. 1509–1523, 2015.
- [7] R. M. Gray, *Conditional rate-distortion theory*. Inf. Sys. Laboratory, Stanford Electronics Laboratories, 1972.
- [8] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [9] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Inf. Theory*, vol. 22, no. 1, pp. 1–10, 1976.
- [10] A. B. Wagner, B. G. Kelly, and Y. Altuğ, "Distributed rate-distortion with common components," *IEEE Trans. on Inf. Theory*, vol. 57, no. 7, pp. 4035–4057, 2011.
- [11] M. Gastpar, "The Wyner-Ziv problem with multiple sources," *IEEE Trans. on Inf. Theory*, vol. 50, no. 11, pp. 2762–2768, 2004.
- [12] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell System Technical Journal*, vol. 52, no. 7, pp. 1037–1076, 1973.
- [13] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM Journal on Applied Mathematics*, vol. 28, no. 1, pp. 100–113, 1975.
- [14] A. D. Wyner, "The common information of two dependent random variables," *IEEE Trans. on Inf. Theory*, vol. 21, no. 2, pp. 163–179, 1975.