Reliable Cooperative Source Transmission with Side Information

Deniz Gündüz, Elza Erkip
Department of Electrical and Computer Engineering,
Polytechnic University,
Brooklyn, NY
dgundu01@utopia.poly.edu, elza@poly.edu

***Abstract*— We consider reliable transmission of a discrete memoryless source over a cooperative relay broadcast channel, where both the relay and the destination terminals want to reconstruct the source; and over a relay channel, where only the destination terminal wishes to obtain a lossless reconstruction. We assume that both the relay and the destination have correlated side information. We find the necessary and sufficient conditions for a general cooperative relay broadcast channel, and for a physically degraded relay channel when the side information at the destination is a degraded version of the relay side information. Our achievability results are based on operational source-channel separation. We utilize source and channel codes that interact only by passing along decoded source codewords from one block to another.**

I. Introduction

We have a limited understanding of general source-channel matching conditions for multiuser networks. Shannon's source-channel separation theorem fails and necessary and sufficient conditions for reliable transmission are not known even for simple scenarios such as transmitting correlated sources over a multiple access channel [1] or over a broadcast channel [2]. There are a limited number of cases such as [3]-[7] for which the lossless source transmission problem is completely solved.

In this paper, we first consider reliable transmission of a discrete memoryless (DM) source over a cooperative relay broadcast (CRB) channel where both the relay and the destination have their own correlated side information. In the CRB scenario, both the relay and the destination terminals want a lossless reconstruction of the source signal. We find the necessary and sufficient conditions for this problem and show that a novel strategy based on superposition block Markov encoding and backward decoding in the joint source-channel coding setting achieves the optimal performance.

Next, we consider transmission of a DM source over a relay channel, where only the destination terminal is interested in reconstructing the source in a lossless fashion, while the relay behaves as a helper based on both its received signal and correlated side information. While the sufficiency conditions established for the CRB case apply for the general relay channel as well, we show that the same conditions are also necessary for a physically degraded relay channel when the source and the destination side information are independent given the relay side information, that is, the destination side information is degraded with respect to the relay side information.

In [5] it is shown that, when broadcasting a common source to multiple receivers, each with its own side information, Shannon's source-channel separation (*informational separation*) fails, and the optimal performance can be achieved by joint source-channel coding. Tuncel [5] proves the optimality of *operational separation* and considers separate source and channel coders that interact at the decoding stage. Our results show that operational separation is also optimal for the more general CRB channel. In our strategy, which is based on block Markov encoding, the source and channel encoders and decoders at each block act independently; however, we have explicit information transfer among the source and channel coders of consecutive blocks. Therefore, even though the strict informational separation in Shannon sense does not hold, our strategy allows the use of individually optimized source and channel codes, thereby achieving modularity in the system design.

In a related work [8], lossless transmission over a relay channel with side information only at the relay terminal is considered. An achievable scheme using joint source-channel coding with list decoding is proposed. In [9], we consider lossy transmission of a Gaussian source over a Gaussian relay channel with correlated relay side information, and propose several achievable schemes together with a lower bound on the average distortion.

The paper is organized as follows: In Section II, we introduce the system model and in Section III we state our main results. In Section IV, we give the proofs of the main theorems. Finally, we discuss our results in Section V and conclude the paper in Section VI.

Throughout the paper, we will denote random variables by capital letters, sample values by the respective lower case letters, and the alphabets by the respective calligraphic letters. The random vector $(X_1, \ldots, X_n)$ will be denoted by X^n, and the complement of a certain element X_i in a vector X^n by $X_i^c \triangleq (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$.

This work is partially supported by NSF grant No. 0430885 and No. 0635177.

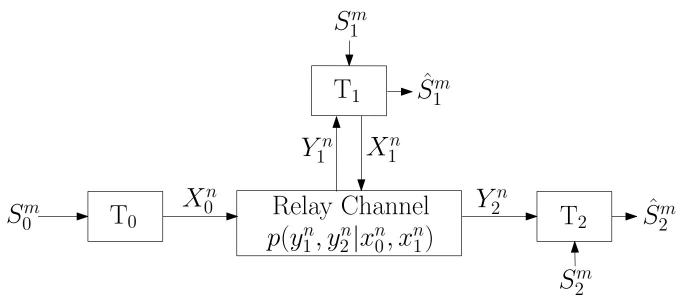

Fig. 1. Relay channel model with node T_0 acting as the source node, node T_2 as the destination node, and T_1 as the relay node.

II. Problem Setup

We have a relay network composed of three terminals (see Fig. 1): $\mathrm{T}_0, \mathrm{T}_1, \mathrm{T}_2$, which are denoted as the source, the relay and the destination node, respectively. The underlying discrete memoryless (DM) relay channel is characterized by the conditional distribution

$$p(y_1^n, y_2^n|x_0^n, x_1^n) = \prod_{i=1}^{n} p_{Y_1,Y_2|X_0,X_1}(y_{1,i}, y_{2,i}|x_{0,i}, x_{1,i}), \quad (1)$$

where X_j, Y_i, $j = 0, 1$, $i = 1, 2$ denote the inputs and outputs at T_j and T_i, respectively.

We consider discrete memoryless i.i.d. sources (S_0, S_1, S_2) which are arbitrarily correlated according to a joint distribution $p(s_0, s_1, s_2) \in \mathcal{P}(\mathcal{S}_0 \times \mathcal{S}_1 \times \mathcal{S}_2)$, that is, $\{S_{0,i}, S_{1,i}, S_{2,i}\}_{i=1}^{\infty}$ are generated i.i.d according to $p(s_0, s_1, s_2)$ over a finite alphabet $\mathcal{S}_0 \times \mathcal{S}_1 \times \mathcal{S}_2$.

We assume, node T_j has access to source S_j, for $j = 0, 1, 2$. In the cooperative relay broadcast scenario, the goal of the source node is to transmit S_0 to both the relay and the destination terminals reliably in Shannon sense. In the relay channel scenario only the destination node needs to reconstruct the source signal losslessly.

Node T_0 maps its observation S_0^m to a channel codeword of length-n by the encoding function $f_0^{(m,n)} : \mathcal{S}^m \to \mathcal{X}_0^n$, i.e., $X_0^n = f_0^{(m,n)}(S_0^m)$. The relay's channel input at each time instant i, denoted as $X_{1,i}$ can depend on the previous channel outputs Y_1^{i-1} and its side information S_1^m. Hence the relay has encoding functions $f_1^{(m,n)} = \{f_{1,1}^{(m,n)}, \dots, f_{1,n}^{(m,n)}\}$ such that

$$X_{1i} = f_{1,i}^{(m,n)}(Y_{1,1}, \dots, Y_{1,i-1}, S_1^m), \qquad 1 \le i \le n.$$

The decoder at node T_i, i=1,2, maps the channel output Y_i^n and its side information S_i^m to the estimate $\hat{S}_i^m$ by the decoding function

$$g_i^{(m,n)} : \mathcal{Y}_i^n \times \mathcal{S}_i^m \to \mathcal{S}_0^m, \quad (2)$$

i.e., $\hat{S}_i^m = g_i^{(m,n)}(Y_i^n, S_i^m)$. Note that, we do not have $g_1^{(m,n)}$ in the relay channel scenario.

The probability of error for the cooperative relay broadcast system is defined as

$$P_e^{(m,n)} = Pr\left[\bigcup_{i=1,2} \{\hat{S}_i^m \neq S_0^m\}\right], \quad (3)$$

while for the relay channel model, we have

$$P_e^{(m,n)} = Pr\left\{\hat{S}_2^m \neq S_0^m\right\}, \quad (4)$$

Definition 2.1: We say that the rate b is *achievable* if, there exist a sequence of encoders $f_j^{(m,n)}$, and decoders $g_j^{(m,n)}$, $j = 1, 2$ with $b = n/m$, such that probability of error vanishes, i.e., $P_e^{(m,n)} \to 0$, as $m, n \to \infty$.

Definition 2.2: A relay channel is said to be *physically degraded* if, $p(y_1, y_2|x_0, x_1)$ can be written in the form

$$p(y_1, y_2|x_0, x_1) = p(y_1|x_0, x_1)p(y_2|y_1, x_1),$$

or, equivalently, if $X_0 \to (X_1, Y_1) \to Y_2$ form a Markov chain for all input distributions $p(x_0, x_1)$.

III. Main Results

This section contains the main results of this paper while the proofs are left to Section IV. Section V contains the discussions of these results. The first theorem states the necessary and sufficient conditions for the CRB channel scenario. In the next theorem we show that the same conditions hold in the case of a degraded relay channel as well, when the sources satisfy a certain Markov chain condition. Finally, we consider a relay channel with feedback.

Theorem 3.1: For the cooperative relay broadcast (CRB) channel with relay and receiver side information as outlined in Section II, rate b is achievable if,

$$H(S_0|S_1) < bI(X_0; Y_1|X_1), \quad (5)$$
$$H(S_0|S_2) < bI(X_0, X_1; Y_2), \quad (6)$$

for some input distribution $p(x_0, x_1)$.

Conversely, if rate b is achievable, then there exists an input distribution $p(x_0, x_1)$ such that (5)-(6) are satisfied with $<$ replaced by $\le$.

Proof: See Section IV-A. ■

The following corollary of Theorem 3.1 was also proved in [5] using a different coding scheme for the direct part.

Corollary 3.2: For lossless broadcasting of a common source S_0 over a broadcast channel to two terminals each with its own correlated side information S_1 and S_2, rate b is achievable if,

$$H(S_0|S_1) < bI(X_0; Y_1), \quad (7)$$
$$H(S_0|S_2) < bI(X_0; Y_2), \quad (8)$$

for some input distribution $p(x_0)$.

Conversely, if rate b is achievable, then there exists an input distribution $p(x_0)$ such that (7)-(8) are satisfied with $<$ replaced by $\le$.

Proof: See Section IV-B. ■

Note that the sufficiency conditions of Theorem 3.1 hold for the relay channel scenario as well. The following theorem proves that the same conditions are also necessary for a degraded relay channel with degraded side information.

Theorem 3.3: For a relay channel with relay and destination side information as outlined in Section II, rate b is achievable if conditions (5)-(6) hold. Conversely, for a degraded relay channel with relay and destination side information satisfying the Markov chain condition $S_0 - S_1 - S_2$, if rate b is achievable, then there exists an input distribution $p(x_0, x_1)$ such that (5)-(6) are satisfied with $<$ replaced by $\leq$.

Proof: See Section IV-C. ■

Finally, for an arbitrary relay channel, we consider perfect feedback from the destination output to the relay terminal.

Theorem 3.4: For an arbitrary relay channel with perfect feedback from the destination channel output to the relay terminal, rate b is achievable if,

$$H(S_0|S_1) < bI(X_0; Y_1, Y_2|X_1), \quad (9)$$

$$H(S_0|S_2) < bI(X_0, X_1; Y_2), \quad (10)$$

for some input distribution $p(x_0, x_1)$. Conversely, assuming that $S_0 - S_1 - S_2$ form a Markov chain for an arbitrary relay channel with perfect destination-relay channel feedback, if rate b is achievable then, (9)-(10) hold with $<$ replaced by $\leq$.

Proof: We can now view the relay output as (Y_1, Y_2), hence we have the Markov chain $X_0 - (X_1, Y_1, Y_2) - Y_2$, and the achievability and converse follow from applying Theorem 3.3. ■

IV. Proofs of the Results

A. Proof of Theorem 3.1

1) Achievability: Fix $p(x_0, x_1)$ such that conditions (5)-(6) hold. We use superposition block Markov encoding, sequential decoding at the relay and backward decoding at the destination.

Code generation: Generate $M = \exp[m(H(S_0) + \epsilon_1)]$ m-length i.i.d. codewords with $\prod_{t=1}^{m} p(s_0)$ and label them as $w(i)$, $i \in [1, M]$. Then randomly partition these codewords into $M_1 = \exp[m(H(S_0|S_1) + \epsilon_2)]$ bins and enumerate these bins with index $w_1 \in [1, M_1]$. We call these bins as the relay bins.

Independent from the relay bins, randomly partition the $w(i)$ codewords into $M_2 = \exp[m(H(S_0|S_2) + \epsilon_3)]$ bins and enumerate these bins with index $w_2 \in [1, M_2]$. We call these as the destination bins.

For the channel codebook, generate M_2 codewords $x_1^n(j)$ for $j \in [1, M_2]$ i.i.d. with $p(x_1^n(j)) = \Pi_{t=1}^{n} p(x_1)$ and index them as $x_1^n(j)$. For each $x_1^n(j)$, generate M_1 conditionally independent codewords $x_0^n(i|j)$, $i \in [1, M_1]$ with probability $p(x_0^n|x_1^n(j)) = \Pi_{t=1}^{n} p(x_0|x_1(j))$.

Encoding: Consider a source sequence S_0^{Bm} of length Bm. Partition this sequence into B portions, $s_{0,b}^m$, $b = 1, \ldots, B$. Similarly, partition the side information sequences into B length-m blocks $s_1^{Bm} = [s_{1,1}^m, \ldots, s_{1,B}^m]$ and $s_2^{Bm} = [s_{2,1}^m, \ldots, s_{2,B}^m]$. We will transmit a total of Bm source samples over a total of $(B+1)n$ channel uses, i.e., over $B+1$ blocks of n channel uses each. For any fixed (m, n) with $n = bm$, we can achieve a rate arbitrarily close to $b = n/m$ by increasing B, i.e., $(B+1)n/Bm \approx n/m = b$.

In block 1, node T_0 observes $s_{0,1}^m$, finds a jointly typical codeword $w(1)$ and the corresponding relay bin index $w_{1,1} \in [1, M_1]$. It transmits the channel codeword $x_0(w_{1,1}, 1)$. The relay simply transmits $x_1(1)$. Similarly, in block b for $b = 2, \ldots, B$, the source terminal transmits the channel codeword $x_0(w_{1,b}, w_{2,b-1})$ where $w_{1,b} \in [1, M_1]$ is the relay bin index of codeword $w(b)$ jointly typical with the source vector $s_{0,b}^m$, and $w_{2,b-1} \in [1, M_2]$ is the destination bin index of the codeword $w(b-1)$ jointly typical with the source vector $s_{0,b-1}^m$. In block $B+1$, node T_0 transmits $x_0(1, w_{2,B})$.

Now, assume that the relay knows $s_{0,b-1}^m$ at the end of block $b-1$. It finds the corresponding destination bin index $\hat{w}_{2,b-1} \in [1, M_2]$. At block b, for $b = 2, \ldots, B+1$, it transmits the channel codeword $x_1(\hat{w}_{2,b-1})$.

Decoding and Error Probability Analysis: The relay decodes the source signal sequentially trying to reconstruct source block $s_{0,b}^m$ at the end of channel block b. Assume that the relay knows $s_{0,b-1}^m$ at the end of block $b-1$ with high probability. Hence, it can find the destination bin index $w_{2,b-1}$. Using this information and its received signal y_1^n, the relay channel decoder will attempt to decode $w_{1,b}$, i.e., the relay bin index corresponding to $s_{0,b}^m$. This is then given to the relay source decoder. With the relay bin index and the side information $s_{1,b}^m$, the relay source decoder estimates $s_{0,b}^m$. The estimation error can be made smaller than ϵ for large enough m, n satisfying $n = bm$ since,

$$H(S_0|S_1) < bI(X_0; Y_1|X_1).$$

Decoding at the destination will be done using backward decoding. The destination node waits till the end of channel block $B+1$. It first tries to decode $s_{0,B}^m$ using the received signal at channel block $B+1$ and its side information $s_{2,B}^m$. Going backwards from the last channel block to the first, we assume that the destination knows $s_{0,b+1}^m$ at the end of channel block $b+2$ for $b = B-1, \ldots, 0$. Hence the destination also knows the relay bin index $w_{1,b+1}$ corresponding to $s_{0,b+1}$. Then at block $b+1$, the destination channel decoder first estimates the bin index $\hat{w}_{2,b}^m$ corresponding to $s_{0,b}^m$ based on its received signal y_2^n. This bin index is then provided to the destination source decoder. The destination source decoder estimates $s_{0,b}^m$ using $\hat{w}_{2,b}^m$ and its side information $s_{2,b}^m$. Arbitrarily small error probability is achieved for the estimate $\hat{s}_{0,b}^m$ at the destination with large enough m, n and $n = bm$ since,

$$H(S_0|S_2) < bI(X_0, X_1; Y_2).$$

2) Converse: Let $P_e^{(m,n)} \to 0$ for a sequence of encoders $f_0^{(m,n)}, f_1^{(m,n)}$ and decoders $g_1^{(m,n)}, g_2^{(m,n)}$ with $n = bm$.

We will use Fano's inequality which states that

$$\begin{aligned} H(S_0^m|\hat{S}_i^m) &\leq 1 + mPr\{S_0^m \neq \hat{S}_i^m\} \log(|\mathcal{S}_0|), \\ &\leq 1 + mP_e^{(m,n)} \log(|\mathcal{S}_0|), \\ &\triangleq m\delta(P_e^{(m,n)}), \end{aligned} \quad (11)$$

where $\delta(x)$ is a nonnegative function that goes to zero as $x \to 0$. Then we can obtain the following chain of inequalities.

$$\sum_{i=1}^{n} I(X_{0,i};Y_{1,i}|X_{1,i}) \geq \sum_{i=1}^{n} I(S_0^m;Y_{1,i}|X_{1,i}), \quad (12)$$
$$= \sum_{i=1}^{n} I(S_0^m, S_1^m;Y_{1,i}|X_{1,i}), \quad (13)$$
$$\geq \sum_{i=1}^{n} I(S_0^m;Y_{1,i}|X_{1,i}, S_1^m), \quad (14)$$
$$= \sum_{i=1}^{n} I(S_0^m, Y_1^{i-1};Y_{1,i}|X_{1,i}, S_1^m) - I(Y_1^{i-1};Y_{1,i}|X_{1,i}, S_0^m, S_1^m), \quad (15)$$
$$= \sum_{i=1}^{n} I(S_0^m, Y_1^{i-1};Y_{1,i}|X_{1,i}, S_1^m), \quad (16)$$
$$\geq \sum_{i=1}^{n} I(S_0^m;Y_{1,i}|Y_1^{i-1}, X_{1,i}, S_1^m), \quad (17)$$
$$= \sum_{i=1}^{n} I(S_0^m;Y_{1,i}|Y_1^{i-1}, S_1^m), \quad (18)$$
$$= I(S_0^m;Y_1^n|S_1^m), \quad (19)$$
$$= H(S_0^m|S_1^m) - H(S_0^m|Y_1^n, S_1^m),$$
$$\geq mH(S_0|S_1) - H(S_0^m|\hat{S}_1^m), \quad (20)$$
$$\geq mH(S_0|S_1) - m\delta(P_e^{(m,n)}), \quad (21)$$

where (12) follows from data processing inequality and the fact that $S_0^m - X_{0,i} - Y_{1,i}$ forms a Markov chain given X_1; (13) follows from the Markov chain $S_1^m - (S_0^m, X_{1,i}) - Y_{1,i}$; (14) follows from the chain rule and nonnegativity of the mutual information; (15) follows from the chain rule; (16) follows from the memoryless channel assumption; (17) again follows from the chain rule and nonnegativity of the mutual information; (18) follows from the fact that $X_{1,i}$ is a function of Y_1^{i-1} and S_1^m; (19) follows from the chain rule; (20) follows from the i.i.d source assumption and the fact that $\hat{S}_1^m$ is a function of Y_1^n and S_1^m; and finally (21) follows from Fano's inequality.

We can also obtain the following chain of inequalities.

$$\sum_{i=1}^{n} I(X_{0,i}, X_{1,i};Y_{1,i}) \geq \sum_{i=1}^{n} I(S_0^m, S_1^m;Y_{2,i}), \quad (22)$$
$$= \sum_{i=1}^{n} I(S_0^m, S_1^m, S_2^m;Y_{2,i}), \quad (23)$$
$$\geq \sum_{i=1}^{n} I(S_0^m, S_1^m;Y_{2,i}|S_2^m), \quad (24)$$
$$= \sum_{i=1}^{n} I(S_0^m, S_1^m, Y_2^{i-1};Y_{2,i}|S_2^m) - I(Y_2^{i-1};Y_{2,i}|S_0^m, S_1^m, S_2^m), \quad (25)$$
$$= \sum_{i=1}^{n} I(S_0^m, S_1^m, Y_2^{i-1};Y_{2,i}|S_2^m), \quad (26)$$
$$\geq \sum_{i=1}^{n} I(S_0^m, S_1^m;Y_{2,i}|Y_2^{i-1}, S_2^m), \quad (27)$$
$$= I(S_0^m, S_1^m;Y_2^n|S_2^m), \quad (28)$$
$$\geq I(S_0^m;Y_2^n|S_2^m), \quad (29)$$
$$\geq H(S_0^m|S_2^m) - H(S_0^m|Y_2^n, S_2^m), \quad (30)$$
$$\geq mH(S_0|S_2) - H(S_0^m|\hat{S}_2^m), \quad (31)$$
$$\geq mH(S_0|S_2) - m\delta(P_e^{(m,n)}), \quad (32)$$

where (22) follows from data processing inequality and the fact that $(S_0^m, S_1^m) - (X_{0,i}, X_{1,i}) - Y_{2,i}$ forms a Markov chain; (23) follows from the Markov chain $S_2^m - (S_0^m, S_1^m) - Y_{2,i}$; (24) follows from the chain rule and nonnegativity of the mutual information; (25) follows from the chain rule; (26) follows from the memoryless channel assumption; (27)-(29) again follow from the chain rule and nonnegativity of the mutual information; (31) follows from data processing inequality since $\hat{S}_2^m$ is a function of Y_2^n and S_2^m; (32) follows from Fano's inequality.

From (21), (32), we have

$$\frac{1}{n}\sum_{i=1}^{n} I(X_{0,i};Y_{1,i}|X_{1,i}) \geq \frac{1}{b}\left(H(S_0|S_1) - \epsilon\right),$$
$$\frac{1}{n}\sum_{i=1}^{n} I(X_{0,i}, X_{1,i};Y_{1,i}) \geq \frac{1}{b}\left(H(S_0|S_2) - \epsilon\right),$$

for any $\epsilon > 0$ and large enough n. We conclude the proof by letting $P_e^{(m,n)} \to 0$, and using the concavity of mutual information over the set of all joint distributions $p(x_0, x_1)$.

B. Proof of Corollary 3.2

Broadcast channel is a special case of our RBC model when the channel outputs y_1 and y_2 are independent of the channel input x_1. Then the necessary conditions directly follow as we have

$$I(X_0;Y_1|X_1) = I(X_0;Y_1), \quad (33)$$
$$I(X_0, X_1;Y_2) = I(X_0;Y_2). \quad (34)$$

Sufficiency also follows by setting $X_1 = \emptyset$ in the achievability of Theorem 3.1. Note that the achievability of Theorem 3.1 uses superposition block Markov encoding at the source node T_0, and backward decoding at node T_2, a strategy not typically used for broadcast channels.

C. Proof of Theorem 3.3

Achievability is the same as in Theorem 3.1. For the converse (similar to the proof of Theorem 3.1), we have the following:

$$\sum_{i=1}^{n} I(X_{0,i};Y_{1,i},Y_{2,i}|X_{1,i})$$

$$\begin{aligned}
&\geq \sum_{i=1}^{n} I(S_0^m;Y_{1,i},Y_{2,i}|X_{1,i}), && (35)\\
&= \sum_{i=1}^{n} I(S_0^m,S_1^m,S_2^m;Y_{1,i},Y_{2,i}|X_{1,i}), && (36)\\
&= I(S_0^m;Y_1^n,Y_2^n|S_1^m,S_2^m), && (37)\\
&= H(S_0^m|S_1^m,S_2^m) - H(S_0^m|Y_1^n,Y_2^n,S_1^m,S_2^m), &&\\
&\geq H(S_0^m|S_1^m) - H(S_0^m|Y_2^n,S_2^m), && (38)\\
&\geq mH(S_0|S_1) - H(S_0^m|\hat{S}_2^m), && (39)\\
&\geq mH(S_0|S_1) - m\delta(P_e^{(m,n)}), && (40)
\end{aligned}$$

where (35)-(37) are similar to (12)-(19); (38) follows from the Markov chain $S_0^m - S_1^m - S_2^m$ and the fact that conditioning reduces entropy; (39) follows from the i.i.d source assumption and the fact that $\hat{S}_2^m$ is a function of Y_2^n and S_2^m; and (40) follows from Fano's inequality.

Since we have a physically degraded relay channel, we have

$$\sum_{i=1}^{n} I(X_{0,i};Y_{1,i},Y_{2,i}|X_{1,i}) = \sum_{i=1}^{n} I(X_{0,i};Y_{1,i}|X_{1,i}),$$

The second part of the proof is given by (22)-(32). Similar to the proof of Theorem 3.1, using the concavity of mutual information we complete the proof.

V. Discussions

The proof of Theorem 3.1 in Section IV shows that, while each source/channel encoder/decoder acts on its own within each block, there is interaction among the channel and source codes of multiple blocks through message transfer. Hence we can use a concatenation of existing near-optimal source and channel codes while maintaining the required source/channel coder interaction. Note that, this is different than the operational separation technique proposed in [5], in which either the decoding is done jointly or the channel decoder outputs a list for the source decoder, which in turn finds the right source message using this list and the side information.

We note that, despite its modular structure, our coding scheme does not provide informational separation in Shannon sense. In informational separation, the source and channel coding is completely separated once the rate is picked. In our scheme, however, the source and channel codes, when considered over B blocks, are not standalone. The relay channel decoder uses the information provided by the relay source decoder from the previous block, similarly the destination channel decoder uses the output of destination source decoder from the next block. Also, the relay uses the decoded source codeword from the previous block to find its transmitted channel codeword.

Our strategy relies on backward decoding which increases the decoding delay at the destination. For relay channel coding, *sliding-window decoding* scheme is known to achieve the same performance as backward decoding with much smaller end-to-end decoding delay. However, a closer examination of our scenario reveals that a straightforward application of sliding-window decoding at the destination would not work in this joint source-channel coding scenario. The channel codes of the source terminal at block b and the relay terminal at block $b+1$ do not correspond to the same message index, and thus decoding them jointly at the destination would perform poorly. A list decoding approach would help reduce the delay with an increased decoding complexity.

Finally, for a general relay channel, a destination-relay channel feedback not only allows us to characterize the necessary and sufficient conditions for lossless transmission in the case of degraded side information, but also leads to the optimality of operational separation.

VI. Conclusion

In this paper, we consider transmission of a discrete memoryless source over cooperative relay broadcast and relay channels where the relay and the destination terminals have access to correlated side information. We prove the optimality of operational separation for an arbitrary cooperative relay broadcast channel, a physically degraded relay channel with degraded side information and finally for an arbitrary relay channel with perfect destination relay feedback and degraded side information. For the achievability, we propose a novel superposition block Markov encoding and backward decoding scheme, which provides modularity in the system design.

Our results can be extended to more complicated relay network scenarios, and can help us design optimal network components to best utilize the available side information at the relay and the destination terminals for joint source-channel coding over cooperative relay (broadcast) network.

References

[1] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," IEEE Trans. Inform. Theory, vol. IT-26, pp. 648-657, Nov. 1980.

[2] Te Sun Han, Max H.M. Costa, "Broadcast channels with arbitrarily correlated sources," IEEE Trans. on Info. Theory, vol. 33, no:5, pp.641-650, 1987.

[3] T. S. Han, "Slepian-Wolf-Cover theorem for a network of channels," Inform. Contr., vol. 47, no. 1, pp. 6783, 1980.

[4] K. de Bruyn, V. V. Prelov, E. van der Meulen, "Reliable transmission of two correlated sources over an asymmetric multiple-access channel," IEEE Trans. on Information Theory, vol. 33, pp. 716 - 718, Sep. 1987.

[5] E. Tuncel, "Slepian-Wolf coding over broadcast channels," IEEE Transactions on Information Theory, vol. 52, no. 4, pp.1469-1482, April 2006.

[6] D. Gündüz, E. Erkip, "Lossless transmission of correlated sources over a multiple access channel with side information," in Proc. of Data Compression Conference (DCC), Salt Lake City, UT, March 2007.

[7] D. Gündüz, E. Erkip, "Interference channel and compound MAC with correlated sources and receiver side information," *IEEE International Symposium on Information Theory (ISIT 2007)*, Nice, France, June 2007.

[8] B. Smith, S. Vishwanath, "Capacity analysis of the relay channel with correlated sources," submitted to IEEE Trans. on Information Theory.

[9] D. Gündüz, C. T. K. Ng, E. Erkip, A. J. Goldsmith, "Source transmission over relay channel with correlated relay side information," IEEE Inter. Symposium on Inform. Theory (ISIT 2007), Nice, France, June 2007.