Duality and the Value of Cooperation in Distributive Source and Channel Coding Problems*

Wei Yu
Electrical and Computer Engineering Department
University of Toronto
10 King's College Road, Toronto, Ontario M5S 3G4, Canada
email: weiyu@comm.utoronto.ca

Abstract

This paper investigates the condition for duality between the achievable rate regions of the indirect distributive source coding problem and the broadcast channel. It has been shown previously that crucial to the existence of duality are two simultaneous Markov conditions that the joint distribution must satisfy. In this paper, we illustrate that the Markov conditions are satisfied if and only if distributive coding achieves the same sum rate as centralized coding under the same joint distribution. Thus, duality is closely related to the value of cooperation. Duality exists if and only if cooperation does not help. Several classes of channels in which duality exists are illustrated as examples. They include a generalization of a previous result on Gaussian multi-terminal source coding and a novel discrete memoryless broadcast channel for which Marton's achievable rate region is optimal at the sum rate point.

1 Introduction

For point-to-point discrete memoryless channels and discrete memoryless sources, it has been known since Shannon that channel coding and source coding can be thought of as duals of each other. The channel capacity problem is a maximization of a mutual information between the channel input and the channel output. The rate-distortion problem is a minimization of a mutual information between the source and its reconstruction. With the identification of a proper test channel, the two mutual information expressions are the same.

Does the duality between source and channel coding extend beyond the point-to-point setting? Toward this direction, Cover and Chiang [1] illustrated that the mutual information expressions for channel coding with side information at the encoder and source coding with side information at the decoder are duals of each other. However, as Pradhan and Ramchandran [2] pointed out, duality truly exists only when the underlying joint distributions of the two problems are identical, which happens only if two sets of Markov chains are simultaneously satisfied. In [3], Pradhan and Ramchandran further generalized the idea to distributive source coding and broadcast channel coding problems. In distributive source coding, two source encoders describe two sources separately and a joint decoder attempts to reconstruct both sources at the receiver. In a broadcast channel, a single transmitter sends multiple messages to separate receivers at the same time. Pradhan and Ramchandran showed that the sum rate expression for the distributive source coding problem is the dual of the sum rate expression

*This research is supported by the Natural Science and Engineering Research Council (NSERC) of Canada under the Canada Research Chair program and through a discovery grant.

for the broadcast channel. Again, duality exists whenever the two underlying probability distributions satisfy two Markov chains at the same time.

The distributive source coding problem described in [3] is often referred to as the "direct coding" problem [4] in the study of so-called "CEO" problem [5] [6] [7] [8]. In [9], Viswanath considered duality for the "indirect coding" problem, in which two correlated observations of a single source are encoded by two encoders separately. While restricting to the Gaussian case, Viswanath showed that the crucial simultaneously Markov chain condition is implied by the recent results on Gaussian vector broadcast channel sum capacity [10] [11] [12]. Thus, corresponding to each Gaussian vector broadcast test channel, there exists a vector Gaussian CEO problems for which the solution can be readily derived.

This paper further studies the "indirect coding" problem and generalizes Viswanath's result to discrete memoryless broadcast channels. First, a duality between the achievable rate region for indirect distributive coding of a general source and Marton's region for a general broadcast channel is illustrated under a simultaneous Markov condition, similar to the approach taken in [3] [9]. The rate region for the source coding problem is a generalization of the earlier work by Berger and Tung [13] [14]. In addition, the simultaneous Markov condition is shown to be closely related to the value of cooperation in distributive coding. Duality exists if and only if distributive source (or channel) coding achieves the same sum rate as joint source (or channel) coding. Finally, several classes of channels for which duality holds are illustrated as examples. They include a generalization of Gaussian distributive source coding problem studied by Viswanath [9] and a novel discrete memoryless broadcast channel for which distributed receivers achieve the same sum rate as a centralized receiver. This broadcast channel gives another example for which Marton's region is optimal at the sum rate point.

2 Distributive Source and Channel Coding

2.1 The CEO Problem

Consider the distributive source coding problem (or the CEO problem [5]) illustrated in Fig. 1. In this problem, Y_1 and Y_2 are two correlated measurements of an underlying random variable X, and they are encoded separately by two different encoders under rate constraints R_1 and R_2, respectively. A joint decoder is interested in reconstructing X with the least amount of distortion possible. This problem is sometime called the "indirect coding" problem in the CEO literature [4]. The general rate-distortion region is still unsolved. In this section, we present an achievable rate-distortion region for this problem.

The achievable region is based on a result by Berger and Tung [13] [14] for the "direct coding" problem. In "direct coding", Y_1 and Y_2 are correlated random variables to be encoded separately, and the joint decoder is interested in reconstructing both Y_1 and Y_2 at the same time. The achievable region for the "indirect" problem differs from the "direct" problem only in the distortion constraint.

Theorem 1 (Berger-Tung for Indirect Coding) *A distortion-rate inner bound for the indirect coding problem at a distortion level D is the set of all (R_1, R_2) such that there exists a pair of random variables (U_1, U_2) with a joint distribution of the form $p(x)p(y_1, y_2|x)p(u_1|y_1)p(u_2|y_2)$ for which*

$$R_1 \geq I(U_1; Y_1) - I(U_1; U_2); \tag{1}$$

$$R_2 \geq I(U_2; Y_2) - I(U_1; U_2); \tag{2}$$

$$R_1 + R_2 \geq I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2), \tag{3}$$

and a deterministic function $\hat{X} = f(U_1, U_2)$ for which $\mathbb{E}[d(X, \hat{X})] \leq D$. Here, X is the

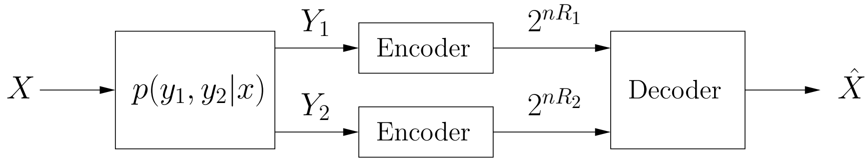

Figure 1: The indirect distributive source coding problem

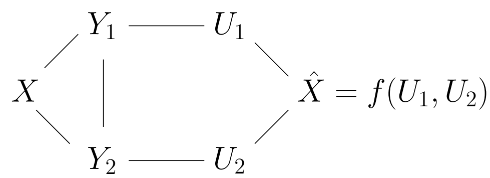

Figure 2: Markov condition for distributive source coding

source to be estimated, $\hat{X}$ is the reconstruction value, Y_1 and Y_2 are observations to be encoded separately, and $d(\cdot, \cdot)$ is the distortion measure.

The proof of this theorem is via a random binning argument and via a Markov Lemma that ensures that the joint typicality of (U_1, U_2, Y_1, Y_2) can be derived from the Markov condition $U_1 - Y_1 - Y_2 - U_2$. The Markov condition is illustrated in Fig. 2 and is crucial for the achievability result to hold. In fact, given the Markov condition, the region simplifies to the following equivalent form:

$$
\begin{aligned}
R_1 &\geq I(U_1; Y_1|U_2); && (4)\\
R_2 &\geq I(U_2; Y_2|U_1); && (5)\\
R_1 + R_2 &\geq I(U_1, U_2; Y_1, Y_2); && (6)
\end{aligned}
$$

This is the form that appeared in Berger and Tung's original work.

2.2 Broadcast Channel

We are motivated to study the "indirect" version of the distributive source coding problem because there is a clear duality between the "indirect" coding problem and the broadcast channel. In a broadcast channel, a joint encoder transmits independent information to multiple uncooperative receivers at the same time. (See Fig. 3.) The capacity region for the general broadcast channel is still an open problem. The best achievable region is due to Marton. Let $p(y_1, y_2|\hat{x})$ be the channel transition probability for the broadcast channel. Marton's region is the set of (R_1, R_2) for which

$$
\begin{aligned}
R_1 &\leq I(U_1; Y_1); && (7)\\
R_2 &\leq I(U_2; Y_2); && (8)\\
R_1 + R_2 &\leq I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2); && (9)
\end{aligned}
$$

over a joint distribution $p(u_1, u_2)p(\hat{x}|u_1, u_2)p(y_1, y_2|\hat{x})$. It is possible to show that the optimal $p(u_1, u_2, \hat{x})$ always gives rise to a deterministic function $\hat{X} = f(U_1, U_2)$. Thus, the random variables involved can be represented graphically in Fig. 4. In particular, an arbitrary intermediate random variable X can be inserted between $\hat{X}$ and (Y_1, Y_2) as long as the Markov condition is preserved.

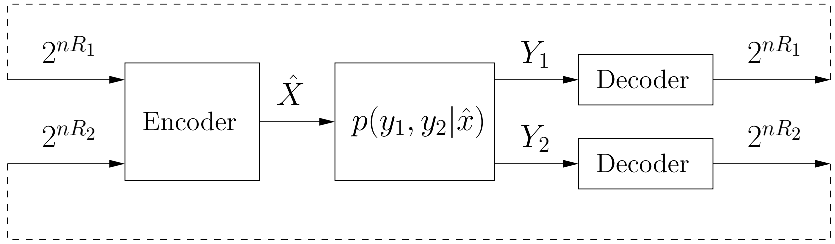

Figure 3: Broadcast channel

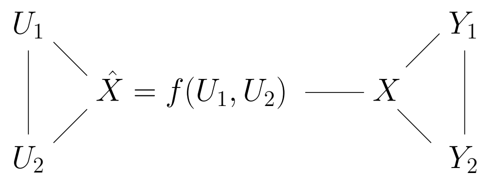

Figure 4: Markov condition for the broadcast channel

3 Duality

There is a striking similarity in the achievable rate region expressions of the distributive source coding and the broadcast channel coding problems, as shown in Fig. 5. The relation between the two coding structures is also evident an inspection of Figs. 1 and 3. The source coding structure is essentially a cyclic shift of the channel coding structure. The encoders of the source coding problem are exactly the decoders of the broadcast channel problem, and the decoder for source coding is exactly the encoder of the broadcast channel. However, for a duality of coding theorems to truly exist, the underlying joint distributions in the two cases would also have to be identical. The identification of the random variables is possible only if the joint distribution satisfies both underlying Markov chain conditions simultaneously. When this happens, a distributive source coding problem can be formulated based on the dual broadcast channel by setting a distortion constraint $D = \mathbb{E}[d(X, \hat{X})]$. Conversely, a broadcast channel problem can be formulated based on the distributive source coding problem by setting the appropriate channel probability $p(y_1, y_2|\hat{x})$. Then, these two problems may be called the duals. The main result of this section is the identification of a condition under which duality holds. It turns out that duality is closely related to the value of cooperation.

Definition 1 *If a joint distribution $p(x, y_1, y_2, u_1, u_2)$ and a deterministic function $\hat{X} = f(U_1, U_2)$ factors in two different ways: $p(x)p(y_1, y_2|x)p(u_1|y_1)p(u_2|y_2)p(\hat{x}|u_1, u_2)$ and $p(u_1, u_2)p(\hat{x}|u_1, u_2)$ $p(y_1, y_2|\hat{x})p(x|\hat{x})$, then their respective distributive source coding and channel coding problems are called duals of each other.*

Theorem 2 *If the distributive source coding problem and the distributive channel coding problem are duals of each other, then in both problems the sum rate for joint coding is the same as the sum rate for separate coding.*

Proof: First, let's consider a broadcast channel for which the joint distribution satisfies the two Markov chain conditions simultaneously. From the source coding Markov chain (Fig. 2):

$$I(\hat{X}; Y_1, Y_2) \leq I(U_1, U_2; Y_1, Y_2). \tag{10}$$

But, by the channel coding Markov chain condition (Fig. 4):

$$I(\hat{X}; Y_1, Y_2) \geq I(U_1, U_2; Y_1, Y_2). \tag{11}$$

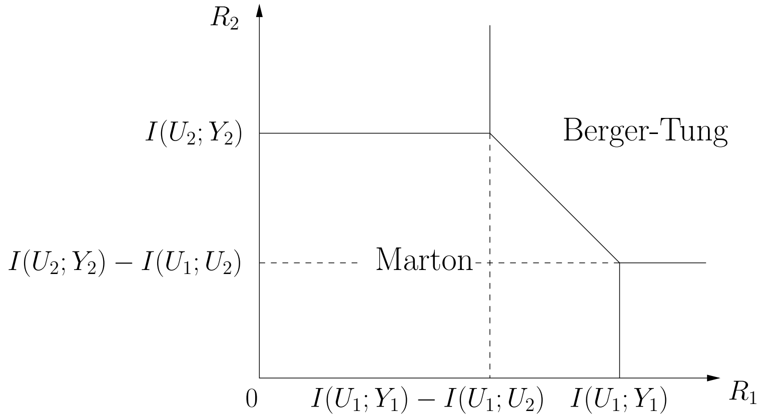

Figure 5: Duality of Marton and Berger-Tung achievable regions

For both to be satisfied, equality must hold. Now, in the broadcast channel, $I(\hat{X};Y_1,Y_2)$ is the cooperative upper bound for the sum rate. On the other hand, Marton's rate region states that $I(U_1;Y_1) + I(U_2;Y_2) - I(U_1;U_2)$ is achievable. But, by the source coding Markov chain, $I(U_1;Y_1) + I(U_2;Y_2) - I(U_1;U_2) = I(U_1,U_2;Y_1,Y_2)$. As $I(\hat{X};Y_1,Y_2) = I(U_1,U_2;Y_1,Y_2)$, this implies that Marton's achievable sum rate reaches the cooperative upper bound, and therefore is optimal.

Conversely, consider the distributive source coding problem for which the joint distribution $p(x, y_1, y_2, u_1, u_2, \hat{x})$ satisfies the two Markov chains simultaneously. From the source coding Markov condition (Fig. 2), we have:

$$I(\hat{X};X) \geq I(U_1,U_2;Y_1,Y_2). \tag{12}$$

But by the channel coding Markov chain condition (Fig. 4):

$$I(\hat{X};X) \leq I(U_1,U_2;Y_1,Y_2). \tag{13}$$

For both to be satisfied, equality must hold. Now, for the distributive source coding problem, $I(\hat{X};X)$ is a cooperative lower bound. On the other hand, Berger-Tung's result states that $I(U_1,U_2;Y_1,Y_2)$ is achievable. Since the two are equal, this means that Berger-Tung's sum rate reaches the cooperative lower bound and is therefore optimal. □

Lemma 1 *Markov chains* $(X_1, X_2) - X_3 - X_4$ *and* $X_1 - X_2 - X_3$ *imply* $X_1 - X_2 - X_3 - X_4$.

Theorem 3 (Viswanath) *If the sum rate of a broadcast channel achieves its cooperative upper bound, then its joint distribution* $p(\hat{x}, y_1, y_2, u_1, u_2)$ *induces a dual distributive source coding problem in which the sum rate achieves the cooperative lower bound.*

Proof: This theorem is essentially the same as an argument made in [9] for the Gaussian channel. The proof is presented here for completeness. Consider a broadcast channel with the associated channel coding Markov chain, where $\hat{X}$ is the input and Y_1 Y_2 are outputs. If a joint receiver achieves the same rate as two separate receivers, then:

$$I(\hat{X};Y_1,Y_2) = I(U_1;Y_1) + I(U_2;Y_2) - I(U_1;U_2). \tag{14}$$

From the channel coding Markov chain, the following chain of inequalities must hold:

$$\begin{aligned} I(\hat{X};Y_1,Y_2) &\geq I(U_1,U_2;Y_1,Y_2) && (15)\\ &= I(U_1;Y_1,Y_2) + I(U_2;Y_1,Y_2|U_1) && (16)\\ &= I(U_1;Y_1,Y_2) + I(U_2;Y_1,Y_2,U_1) - I(U_1;U_2) && (17)\\ &\geq I(U_1;Y_1) + I(U_2;Y_2) - I(U_1;U_2) && (18) \end{aligned}$$

Combining the above two results, it is clear that equalities must hold in (18). For this to happen, the markov chains $(U_1, Y_1) - Y_2 - U_2$ and $U_1 - Y_1 - Y_2$ must hold. By Lemma 1, the two Markov chains imply $U_1 - Y_1 - Y_2 - U_2$. This is exactly the Markov chain condition needed in source coding. In particular, if X is chosen to be an intermediate step in $\hat{X} - X - (Y_1, Y_2)$, then a source coding problem for the random variable X can be formulated with the distortion constraint $\mathbb{E}[d(X, \hat{X})] = D$. Since the variable X is floating and the distortion constraint is arbitrary, the source coding problem is not unique. Finally, as the dual source coding problem automatically satisfies both Markov chains, the dual problem also achieves the cooperative lower bound. □

Conversely, it is also possible to derive a dual broadcast channel from a source coding problem in which cooperative bound is achieved. The following elementary lemma is helpful in proving the result.

Lemma 2 *If random variables $X - Y - Z$ forms a Markov chain and further $I(X;Z) = I(Y;Z)$, then $Y - X - Z$ also forms a Markov chain.*

Theorem 4 *If the sum rate of a distributive source coding problem achieves its cooperative lower bound, then its joint distribution $p(\hat{x}, y_1, y_2, u_1, u_2)$ induces a dual broadcast channel coding problem whose sum rate achieves its cooperative upper bound.*

Proof: Since the sum rate of the distributive source coding achieves its cooperative lower bound, then

$$I(\hat{X}; X) = I(U_1, U_2; Y_1, Y_2). \tag{19}$$

By the source coding Markov chain $X - (Y_1, Y_2) - (U_1, U_2) - \hat{X}$, this implies $I(X, Y_1, Y_2; U_1, U_2) = I(X, Y_1, Y_2; \hat{X})$. By Lemma 2, this implies $(Y_1, Y_2, X) - \hat{X} - (U_1, U_2)$. Similarly, a symmetric argument gives $(Y_1, Y_2) - X - (\hat{X}, U_1, U_2)$. Thus, $(Y_1, Y_2) - X - \hat{X}$. By Lemma 1, the two Markov chains imply

$$(U_1, U_2) - \hat{X} - X - (Y_1, Y_2). \tag{20}$$

This is exactly the channel coding Markov chain. Thus, there is a dual broadcast channel problem associated with the joint distribution $p(u_1, u_2, x, y_1, y_2)$ and the deterministic function $\hat{X} = f(U_1, U_2)$, with $\hat{X}$ as the input and (Y_1, Y_2) as the outputs. Further, since this joint distribution satisfies the two Markov chains simultaneously, the dual broadcast channel problem achieves the cooperative sum rate bound. □

To summarize, if a joint distribution satisfies both the distributive source coding Markov chain and the broadcast channel Markov chain, then it induces a pair of dual source coding and channel coding problems. Further, duality implies that the minimum rate for distributive source coding is exactly $I(X; \hat{X})$, and the maximum rate for the broadcast channel is exactly $I(X; Y_1, Y_2)$. Conversely, if the sum rate of a broadcast channel achieves the cooperative upper bound, or if the sum rate of a distributive source coding achieves the cooperative lower bound, then duality exists. Thus, duality exists if and only if separate coding performs as well as joint coding. The duality of the achievable rate regions for the distributive source coding problem and for the broadcast channel is illustrated in Fig. 5.

Note that in this paper, duality is interpreted for each fixed input distribution. We have deliberately avoided input optimization for the following reason. Consider the maximization of Marton's achievable sum rate $I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2)$ over $p(u_1, u_2, x)$. Suppose that the optimal distribution achieves $I(X; Y_1, Y_2)$. Then, duality holds. However, this does not necessarily imply that

$$\max_{p(u_1, u_2, x)} I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2) = \max_{p(x)} I(X; Y_1, Y_2), \tag{21}$$

as the maximization of $I(X;Y_1,Y_2)$ may yield a different $p(x)$. Thus, the set of problems in which joint coding achieves the same rate as separate coding under a fixed distribution is larger than the set of problems in which joint coding achieves the same rate as separate coding under their respective optimal distribution.

Finally, we caution that both the Marton's region and the Berger-Tung's region are only achievable regions still. Converses have not been established in either case.

4 Examples

4.1 Deterministic Channel

Cover [15] pointed out that there is a duality between the deterministic broadcast channel and the Slepian-Wolf problem. This duality can be interpreted in the current context as follows. Consider a broadcast channel with deterministic outputs $Y_1 = f_1(X)$ and $Y_2 = f_2(X)$. Set $U_1 = Y_1$ and $U_2 = Y_2$. It is easy to verify that the sum rate achieves the cooperative bound: $R_1 + R_2 = H(Y_1, Y_2) = I(X; Y_1, Y_2)$. In fact, both Markov chains are trivially satisfied. Thus, there is duality. The dual source coding problem is the Slepian-Wolf [16] problem. Here, the distortion constraint is zero, i.e. $\hat{X} = X$, and the minimum sum rate is precisely $H(Y_1, Y_2)$.

4.2 Gaussian Vector Channel

We now interpret and give a generalization of Viswanath's result for Gaussian multi-terminal source coding [9]. For duality to hold, there must be no rate loss in the broadcast channel problem. This is true if and only if the noise distribution is the least favorable [10] [11] [12]. Now, the noise in the broadcast channel is the sum of the noise induced by $p(y_1, y_2|x)$ and the distortion in the dual distributive source coding problem. Thus, for duality to hold, the sum must be a least favorable noise.

More concretely, for a Gaussian vector broadcast channel $y = Hx+z$ under a fixed Gaussian input distribution $\mathcal{N}(0, S_{xx})$. The worst noise can be found via a dual multiple access channel with an appropriate input constraint. Let Ψ be the input covariance and λ be the water-filling level in the dual channel. The worst noise is [17]:

$$S_{zz} = H(H^T \Psi H + \lambda I)^{-1} H^T + S'_{zz}, \tag{22}$$

where S'_{zz} is an arbitrary positive semi-definite matrix that makes the diagonals of S_{zz} 1. Our duality result implies that S_{zz} may be split into

$$S_{zz} = S_{nn} + H S_{ee} H^T, \tag{23}$$

where S_{nn} is the sensor noise and S_{ee} is the actual distortion in the distributive source coding problem, and a dual distributive source coding problem may be constructed. This is a generalization of Viswanath's approach in which S_{xx} is set to I and S_{nn} is set zero. To construct a realistic source coding problem with S_{ee} as the actual distortion, a specific distortion constraint needs to be imposed.

4.3 Write-Once Channel

Finally, we construct a discrete memoryless broadcast channel for which duality holds. The construction is based on channels with non-causal side information available at the transmitter, in which the capacity with transmitter side information only is the same as if side information is also available at the receiver, i.e. $I(U;Y) - I(U;S) = I(X;Y|S)$. The memory-with-defect channel [18] is one such example. Consider a memory chip in which occasional bits are either

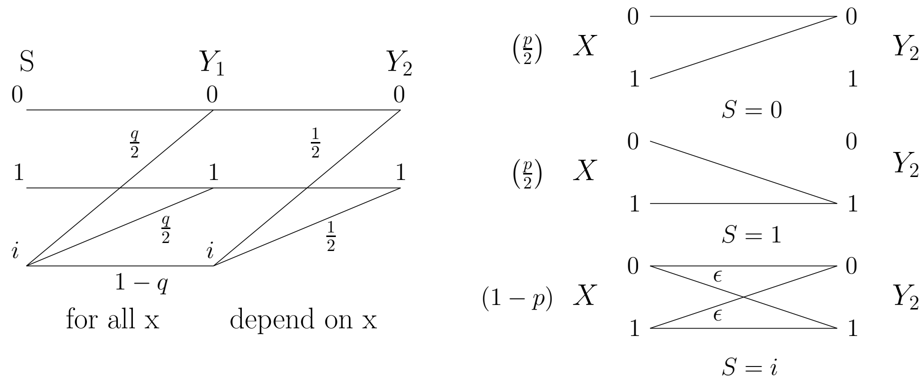

Figure 6: Write-once broadcast channel

stuck at 0 or stuck at 1, and where the location of stuck-at bits are known only at the transmitter. In non-stuck-at locations, the input and output distributions follow a binary symmetric channel with crossover probability ϵ. If the probabilities of stuck-at-0 and stuck-at-1 are equal, then the capacity of the channel with the stuck-at locations known only at the transmitter is the same as if the locations are also known at the receiver [18].

Now, consider a write-once memory. The memory locations are in an initial state i. Writing a bit to a memory location changes its state to either 0 or 1. However, once written, a bit cannot be changed. The transmitter wishes to communicate with two separate receivers using two consecutive memory writes. The first receiver examines the memory after the initial memory write. Then, the transmitter writes to the memory again, after which the second receiver gets a chance to examine the memory. The output alphabet of the first receiver is $\{i, 0, 1\}$. The second receiver, however, can only distinguish between $\{0, 1\}$.

If the first receiver can distinguish among states $\{i, 0, 1\}$ perfectly, then the following strategy seems sensible. The transmitter uses proportion p bits to communicate to the first receiver at a rate $H(p/2, p/2, 1-p)$. Further, additional independent information can be communicated to the second receiver via the defective memory channel at a rate $(1-p)(1-H(\epsilon))$. A careful examination reveals that such a broadcast channel in fact achieves its cooperative upper bound. This can be verified, for example, using known results on broadcast channel with one deterministic component [15].

A more interesting example is when the first receiver sees a corrupted version of the channel state. More precisely, let the input of the write-once channel be (X, S), with an input alphabet $\{0, 1\} \times \{i, 0, 1\}$. The marginal dependence of Y_1 on S is as follows. When $S = 0$ or $S = 1$, then $Y_1 = S$. When $S = i$, then $Y_1 = i$ with probability $1 - q$, $Y_1 = 0$ or 1 with probabilities $q/2$ each. The output Y_2 depends on both X and S as in memory-with-defect. When $S = 0$ or $S = 1$ (the stuck-at states), then $Y_2 = S$. When $S = i$, the distribution of (X, Y_1) follows that of a binary symmetric channel with crossover probability ϵ. For reasons that will be explained later, we require $q \leq 2\epsilon$. Clearly, the transmitter must use S to communicate to Y_1 and X to Y_2. As Y_1 sees a better version of S, and Y_2 sees a better version of X, this broadcast channel is not degraded. Fig. 6 illustrates the write-once channel.

To use the memory-with-defect broadcast strategy, fix the input distribution as follows: let $p(s) = \{p/2, p/2, 1-p\}$ as shown in Fig. 6. Let $p(x|s = i) = \{1/2, 1/2\}$, $p(x|s = 0) = \{1-\epsilon, \epsilon\}$ and $p(x|s = 1) = \{\epsilon, 1-\epsilon\}$. To convey information to Y_1, the transmitter sets p proportion of S to 0 or 1 and keeps the rest at state i. The capacity to Y_1 is $R_1 = I(S; Y_1)$. To convey information to Y_2, the transmitter sets the rest $(1-p)$ proportion of S to 0 or 1, while regarding the previously set states as stuck-at states. Using a memory-with-defect strategy, the capacity to Y_2 is $R_2 = I(U; Y_2) - I(U; S) = I(X; Y_2|S)$. (The optimal U turns out to be X in this case.)

The hope is that by choosing the joint distribution $p(y_1, y_2|x, s)$ judiciously, this broadcast strategy can be shown to achieve the cooperative upper bound. Clearly such a joint distribution has to be a "worst" distribution as in Sato's upper bound [19]. However, for this channel, the worst distribution alone is not yet enough. Consider

$$
\begin{aligned}
I(X,S;Y_1,Y_2) &= I(X,S;Y_1) + I(X,S;Y_2|Y_1) && (24)\\
&= I(S;Y_1) + I(X;Y_1|S) + I(S;Y_2|Y_1) + I(X;Y_2|Y_1,S). && (25)
\end{aligned}
$$

For the above to be equal to $I(S;Y_1)+I(X;Y_2|S)$, three conditions have be met: $I(X;Y_1|S) = 0$, $I(S;Y_2|Y_1) = 0$ and $I(X;Y_2|Y_1,S) = I(X;Y_2|S)$. To construct a joint distribution for which this is true, start with the condition $I(S;Y_2|Y_1) = 0$. We must have the Markov chain $S - Y_1 - Y_2$, as shown in Fig. 6. The trick is to let some of the transitions be dependent on X. In particular, let $p(Y_1 = k|S = i, X = 0) = p(Y_1 = k|S = i, X = 1)$ for $k = \{0, 1, i\}$, so that $I(X;Y_1|S) = 0$. Also, let $p(Y_2 = 0|Y_1 = i, X = 0) = p(Y_2 = 1|Y_1 = i, X = 1) = 1 - \epsilon'$, and $p(Y_2 = 1|Y_1 = i, X = 0) = p(Y_2 = 0|Y_1 = i, X = 1) = \epsilon'$, where ϵ' is chosen to make the marginal distribution $p(y_2|x)$ that of a binary symmetry channel with crossover probability ϵ. This requires

$$\frac{q}{2} + (1-q)\epsilon' = \epsilon, \tag{26}$$

(hence the condition $q \le 2\epsilon$.) Finally, to satisfy the third condition $I(X;Y_2|Y_1,S) = I(X;Y_2|S)$, notice that $I(X;Y_2|Y_1 = 0, S) = I(X;Y_2|Y_1 = 1, S) = 0$ since Y_1 completely determines Y_2. This happens with probability q. With probability $(1-q)$, $I(X;Y_2|Y_1 = i, S) = 1 - H(\epsilon')$. For $I(X;Y_2|Y_1,S)$ to be equal to $I(X;Y_2|S)$, which is $1 - H(\epsilon)$, we require

$$(1-q)(1 - H(\epsilon')) = 1 - H(\epsilon). \tag{27}$$

Thus, $I(X;Y_2|Y_1,S) = I(X;Y_2|S)$ only for a particular value of q that satisfies both (26) and (27). For example, $q = 0.1$ and $\epsilon = 0.198$ is such a choice. With such carefully chosen ϵ, q and joint distribution $p(y_1, y_2|x, s)$, the discrete memoryless broadcast channel can now achieve the same sum rate with separate decoding as with joint decoding under a fixed input distribution $p(x, s)$. Note that the only condition on $p(x, s)$ is that $p(S = 0) = p(S = 1)$ and $p(X = 0|S = i) = p(X = 1|S = i)$. The optimal input distribution that maximizes $I(X,S;Y_1,Y_2)$ clearly takes this form. Therefore, $\max_{p(x,s)} I(X,S;Y_1,Y_2)$ is indeed achievable using Marton's broadcast strategy. However contrived, this is nevertheless a novel example for which Marton's strategy is sum-rate optimal, and it is a non-trivial extension of the Gaussian vector broadcast channel result [10] [11] [12] where a similar proof technique is used.

Finally, we return to duality and illustrate the dual distributive source coding problem for the write-once broadcast channel. Two separate encoders observe Y_1 and Y_2 after the first and the second write, respectively. They individually compress their own observation and report to a joint decoder under a sum rate constraint. The joint decoder is interested in determining (X, S) with the least possible distortion. Since the broadcast channel achieves the cooperative sum rate, the dual distributive source coding problem also achieves the same sum rate with $R_1 = I(S;Y_1)$ and $R_2 = I(X;Y_2|S)$.

5 Conclusion

There exists a duality between the broadcast channel and the distributive source coding problem. However, duality exists if and only if the achievable sum rate is the same with or without cooperation. The deterministic channel, the Gaussian vector channel with the worst-noise, and the write-once channel are illustrated as examples. In general, however, examples of duality are few and far between.

References

[1] T. Cover and M. Chiang, "Duality between channel capacity and rate distortion with two-sided state information," *IEEE Trans. Info. Theory*, pp. 1629–1638, June 2002.

[2] S. Pradhan, J. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Trans. Info. Theory*, May 2003.

[3] S. Pradhan and K. Ramchandran, "On functional duality between MIMO source and channel coding with one-sided collaboration," in *IEEE Info. Theory Workshop*, Oct 2002.

[4] R. Zamir, "Multiterminal source coding with high resolution," *IEEE Trans. Info. Theory*, vol. 45, no. 1, pp. 106–117, Jan 1999.

[5] T. Berger and H. Viswanathan, "The CEO problem," *IEEE Trans. Info. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.

[6] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Info. Theory*, vol. 43, no. 5, pp. 1549–1559, Sep 1997.

[7] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Info. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov 1997.

[8] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Info. Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.

[9] P. Viswanath, "Sum rate of multiterminal Gaussian source coding," in *DIMACS Workshop on Network Information Theory*, March 2003.

[10] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *submitted to IEEE Trans. Info. Theory*, Nov. 2001.

[11] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Info. Theory*, July 2003.

[12] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum-rate capacity of Gaussian MIMO broadcast channels," *submitted to IEEE Trans. Info. Theory*, August 2002.

[13] S.-Y. Tung, *Multiterminal rate-distortion theory*, Ph.D. thesis, Cornell University, 1977.

[14] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications*, G. Longo, Ed. 1977, Springer - Verlag.

[15] T. Cover, "Comments on broadcast channels," *IEEE Trans. Info. Theory*, vol. 44, no. 6, pp. 2524–2530, Oct. 1998.

[16] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Info. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[17] W. Yu, "The structure of the least-favorable noise in Gaussian vector broadcast channels," in *DIMACS Workshop on Network Information Theory*, March 2003.

[18] C. Heegard and A. El Gamal, "On the capacity of computer memories with defects," *IEEE Trans. Info. Theory*, vol. 29, pp. 731–739, Sep. 1983.

[19] H. Sato, "An outer bound on the capacity region of broadcast channels," *IEEE Trans. Info. Theory*, vol. 24, no. 3, pp. 374–377, May 1978.