An Optimal Transmit-Receive Rate Tradeoff in Gray-Wyner Network and Its Relation to Common Information

Kumar Viswanatha, Emrah Akyol and Kenneth Rose
ECE Department, University of California - Santa Barbara
{kumar,eakyol,rose}@ece.ucsb.edu

***Abstract*—This paper considers the problem of characterizing the optimal tradeoff between the total transmit versus receive rate in the Gray-Wyner network. This tradeoff plays a crucial role in many important practical applications including establishing fundamental limits in databases for correlated sources and in minimum cost routing for networks. We develop the insight into this tradeoff by defining two quantities $C(X,Y;R^{'})$ and $K(X,Y;R^{''})$, which quantify the shared rate as a function of the total transmit and receive rates respectively. Closely tied up with this tradeoff is the notion of common information of two dependent random variables. The two most influential definitions are due to Wyner [2] and Gács-Körner [1]. Though it is well know that these definitions can be characterized as two extreme points in the Gray-Wyner region, no contour with operational significance is known which connects them. We will show that the tradeoff between transmit and receive rates leads to a contour of points on the boundary of Gray-Wyner region which passes through the operating points of Wyner and Gács-Körner. We use these properties to derive alternate characterizations for the two definitions of common information under a broader unified framework.**

***Index Terms*—Gray-Wyner network, Common information, Multi-terminal source coding**

I. Motivation and Prior Work

Consider two correlated sources, X and Y, to be compressed and stored in a database for selective retrieval as shown in Figure 1a. Assume that the users query for the two sources individually. A very interesting tradeoff arises in this setup between the total storage rate and the total retrieval rate. At one extreme, we can compress the sources at the minimum storage rate of $H(X,Y)$. However, then the retrieval rate for the individual queries would be more than $H(X)$ and $H(Y)$ respectively. On the other hand, if we compress the sources at their individual entropies, the storage rate entails suboptimality. This paper addresses this tradeoff and its connection with the notions of common information.

This tradeoff not only plays a crucial role in characterizing the performance limits for a database, but also has implications in other important network settings. We recently introduced a new routing paradigm for minimum cost communication of correlated sources over a network called '*dispersive information routing*' (DIR) [9] wherein the intermediate nodes are allowed to '*split*' a packet and forward different subsets of the received bits on each of the forward paths. For example consider one of the simplest network setups shown in Figure 1b, with a single intermediate node, which is allowed to forward different subsets of the received bits on each of the forward parts. DIR can be equivalently realized by the encoder transmitting 3 smaller packets; first packet at rate R_0 destined to both the sinks. Two other packets at rates R_1 and R_2 are meant to the individual sinks as shown in the figure. Here the cost of transmission by the source and the collector depend on $R_0+R_1+R_2$ and $2R_0+R_1+R_2$ respectively. Here again, the transmit-receive rate tradeoff plays a crucial role in determining the transmission rates. This setup is effectively modeled using the Gray-Wyner network shown in Figure 1c.

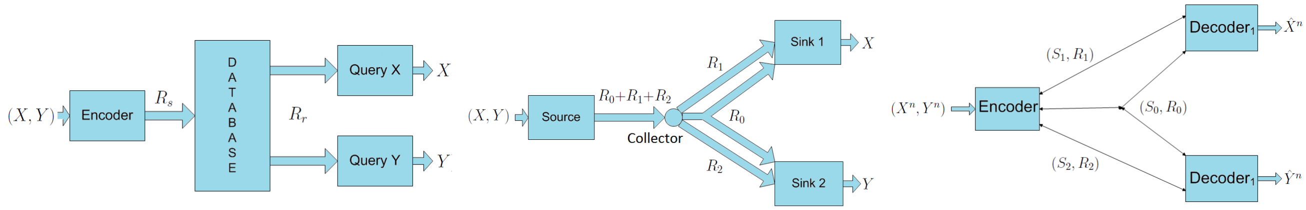
The Gray-Wyner setup [3] has one encoder and two decoders. The encoder observes both X^n and Y^n and generates three descriptions (S_0,S_1,S_2) at respective rates (R_0,R_1,R_2). The first decoder receives (S_0,S_1) and the second receives (S_0,S_2). The two decoders respectively reconstruct X^n and Y^n losslessly. In analogy to a database, S_0 are the set of stored bits which are retrieved for both the queries. S_1 and S_2 are the bits which are individually retrieved for queries X and Y respectively. The complete rate region for the tuple (R_0,R_1,R_2) for loss-less reconstruction of the respective sources at the two decoders is due to Gray and Wyner [3] and is denoted by $\mathcal{R}_{GW}$. For any U jointly distributed with (X,Y), let $\mathcal{R}_{GW}(U)$ be given by:

$$R_0 \geq I(X,Y;U), \; R_1 \geq H(X|U), \; R_2 \geq H(Y|U) \quad (1)$$

then, $\mathcal{R}_{GW} = \bigcup_U \mathcal{R}_{GW}(U)$. We refer to the branch that goes to both the decoders as the 'shared branch' and the other two as 'private branches'.

Note that the total transmit rate is $R_0+R_1+R_2$ and the total receive rate is $2R_0+R_1+R_2$. Our primary objective in this paper is to characterize this tradeoff between transmit rate and receive rate and study its properties. To gain insight into this tradeoff, we characterize two curves which are rotated/transformed versions of each other in Section II. The first curve, denoted by $C(X,Y;R^{'})$, plots the minimum shared rate, R_0, at a transmit rate of $H(X,Y)+R^{'}$ and the second, denoted by $K(X,Y;R^{''})$, is the maximum R_0 (minimum $-R_0$) at a receive rate of $H(X)+H(Y)+R^{''}$. It is easy to see that the transmit-receive rate tradeoff can be derived directly from these quantities. The quantities $C(X,Y;R^{'})$ and $K(X,Y;R^{'})$, while evidently are "transformed" versions of

The work was supported by the NSF under grants CCF-0728986 and CCF - 1016861



(a) Database (b) DIR (c) Gray-Wyner Network

Figure 1: (a) Database illustrating the tradeoff between the storage rate R_s and total retrieval rate R_r (b) A simple example to demonstrate dispersive information routing. (c) The Gray-Wyner network

each other, play important roles in different practical scenarios. For example, $C(X,Y;R^{'})$ characterizes the minimum receive rate at a fixed transmit rate making it more suitable for network applications where the transmit power is constrained. On the other hand, $K(X,Y;R^{''})$ minimizes transmit rate at a fixed receive rate (retrieval rate), which finds applicability in database applications.

Closely related to the understanding of these tradeoffs is the notion of common information of two dependent random variables, which has been actively studied and pursued by researchers over three decades. The two most influential of these definitions are due to Wyner [2] and Gács and Körner [1]. It is well known that the two definitions can be characterized using $\mathcal{R}_{GW}$ and the corresponding operating points are two points on the boundary of the region. Several approaches have been proposed to provide further insight into the underlying connections between them [4], [5], [6], [8]. However, to the best of our knowledge, there has been no work which attempts to find a contour with operational significance, which connects the two operating points on the Gray-Wyner region. In Section III we will show that the contour of points on the boundary obtained by trading-off transmit rate and receive rate passes through both the operating points of Wyner and Gács-Körner. Hence, this tradeoff provides a generic framework to understand the underlying principles of shared information. We note that the quantities $C(X,Y;R^{'})$ and $K(X,Y;R^{''})$ are in fact generalizations of Wyner's and Gács-Körner's definitions to excess sum rate and excess receive rate regimes respectively. Using their properties, we will also derive alternate characterizations for the two notions of common information under a unified framework in Section IV. We note in passing that there have been certain other definitions of common information in the literature [4], [8]. Their relations to our tradeoff curves are less direct and are beyond the scope of this paper.

II. Characterizing $C(X,Y;R^{'})$ and $K(X,Y;R^{''})$

Let (X,Y) be any two dependent random variables taking values in some finite alphabets $\mathcal{X}$ and $\mathcal{Y}$ respectively. We define the quantity $C(X,Y;R^{'})\ \forall R^{'} \in [0, I(X,Y)]$ as:

$$C(X,Y;R^{'}) = \inf R_0 : (R_0,R_1,R_2) \in \mathcal{R}_{GW} \quad (2)$$

satisfying,

$$R_0 + R_1 + R_2 = H(X,Y) + R^{'} \quad (3)$$

Similarly, we define the quantity $K(X,Y;R^{''})\ \forall R^{''} \in [0, H(X,Y) - I(X,Y)]$ as:

$$K(X,Y;R^{''}) = \sup R_0 : (R_0,R_1,R_2) \in \mathcal{R}_{GW} \quad (4)$$

satisfying,

$$2R_0 + R_1 + R_2 = H(X) + H(Y) + R^{''} \quad (5)$$

Note that we restrict the ranges for $R^{'}$ and $R^{''}$ for practical considerations as operating at $R^{'} > I(X;Y)$ or $R^{''} > H(X,Y) - I(X,Y)$ would clearly lead to suboptimality. The following Theorem provides information theoretic characterizations for $C(X,Y;R^{'})$ and $K(X,Y;R^{''})$.

Theorem 1. *(i) For any excess sum rate $R^{'} \in [0, I(X,Y)]$, $C(X,Y;R^{'})$ satisfies:*

$$C(X,Y;R^{'}) = \min I(X,Y;U) \quad (6)$$

where the minimization is over all U jointly distributed with (X,Y) such that:

$$I(X;Y|U) = R^{'} \quad (7)$$

We denote the operating point in $\mathcal{R}_{GW}$ corresponding to the minimum by $P_{C(X,Y)}(R^{'})$.
(ii) For any excess reception rate $R^{''} \in [0, H(X,Y) - I(X,Y)]$, $K(X,Y;R^{''})$ satisfies:

$$K(X,Y;R^{''}) = \max I(X,Y;W) \quad (8)$$

where the maximization is over all W jointly distributed with (X,Y) such that:

$$I(X;W|Y) + I(Y;W|X) = R^{''} \quad (9)$$

We denote the operating point in $\mathcal{R}_{GW}$ corresponding to the maximum by $P_{K(X,Y)}(R^{''})$.

Proof: We prove the theorem only for $C(X,Y;R^{'})$. The proof for $K(X,Y;R^{''})$ follows in similar lines.

Achievability : Say we are given a U jointly distributed with (X,Y) such that $I(X;Y|U) = R^{'}$. It leads to a point in the Gray-Wyner region with $(R_0,R_1,R_2) = (I(X,Y;U), H(X|U), H(Y|U))$. On substituting in (3) we have:

$$\begin{aligned} R_0 + R_1 + R_2 &= I(X,Y;U) + H(X|U) + H(Y|U) \\ &= H(X,Y) + I(X;Y|U) \quad (10) \\ &= H(X,Y) + R^{'} \quad (11) \end{aligned}$$

Note that the existence of a U which achieves the minimum in (6) follows from Theorem 4.4 (A) in [2].

Converse : We know from the converse to the Gray-Wyner region that every point in $\mathcal{R}_{GW}$ is achieved by some random variable U jointly distributed with (X, Y). We need to determine the condition on U for (3) to satisfy. On substituting $(R_0, R_1, R_2) = (I(X, Y; U), H(X|U), H(Y|U))$ in (3), we get the condition to be (7), proving the converse. ■

Note that the cardinality of $\mathcal{U}$ (denoted by $|\mathcal{U}|$) can be restricted to $|\mathcal{U}| \leq |\mathcal{X}| \, . \, |\mathcal{Y}| + 1$ using Theorem 4.4 in [2]. Also note that when the transmit rate is $H(X, Y) + R^{'}$, the minimum receive rate is $H(X, Y) + R^{'} + C(X, Y; R^{'})$. Similarly, when the receive rate is $H(X) + H(Y) + R^{''}$, the minimum transmit rate is $H(X)+H(Y)+R^{''}-K(X, Y; R^{''})$. Hence the quantities $C(X, Y; R^{'})$ and $K(X, Y; R^{''})$ are just rotated/transformed versions of the transmit versus receive rate tradeoff curve.

We refer to plots of $C(X, Y; R^{'})$ and $K(X, Y; R^{''})$ versus $R^{'}$ and $R^{''}$ as the '*transmit tradeoff curve*' and the '*receive tradeoff curve*' respectively. Observe that in both the cases, as we increase $R^{'}$ (or $R^{''}$) we obtain parallel cross-sections of the Gray-Wyner region and the set of operating points $P_{C(X,Y)}(R^{'})$ and $P_{K(X,Y)}(R^{''})$, trace contours on the boundary of the region. We refer to them as the '*transmit contour*' and the '*receive contour*' respectively. Note the difference between the contours and their respective tradeoff curves. The contours are defined in a 3-D space and lie on the boundary of $\mathcal{R}_{GW}$. In general each of the contours may not even lie on a single plane. However, the tradeoff curves are a projection of the respective contours on to a 2-D plane.

III. Relation to Common Information

Let us start with Wyner's common information, even though it is not the earliest, simply because it directly builds on the Gray-Wyner network. It is defined as the minimum rate on the shared branch, while the total sum rate is constrained to be the joint entropy. Formally,

$$C_W(X, Y) = \min R_0 : (R_0, R_1, R_2) \in \mathcal{R}_{GW} \tag{12}$$

subject to,

$$R_0 + R_1 + R_2 = H(X, Y) \tag{13}$$

He showed that $C_W(X, Y) = \inf I(X, Y; U)$, where the infimum is over all random variables U such that $X \leftrightarrow U \leftrightarrow Y$ form a Markov chain in that order. We denote the operating point in $\mathcal{R}_{GW}$ corresponding to the infimum by P_W. It is clear from the definition that the quantity $C(X, Y; R)$ directly generalizes Wyner's idea of common information to the '*excess sum rate*' regime. Observe that $C_W(X, Y) = C(X, Y; 0)$ and $P_{C(X,Y)}(0) = P_W$.

On the other hand, Gács and Körner [1] defined common information (denoted here by $C_{GK}(X, Y)$) as the maximum amount of information relevant to both random variables, which one can extract from the knowledge of either one of them. Formally, they defined common information of X and Y as:

$$C_{GK}(X, Y) = \sup \frac{1}{n} H(f_1(X^n)) \tag{14}$$

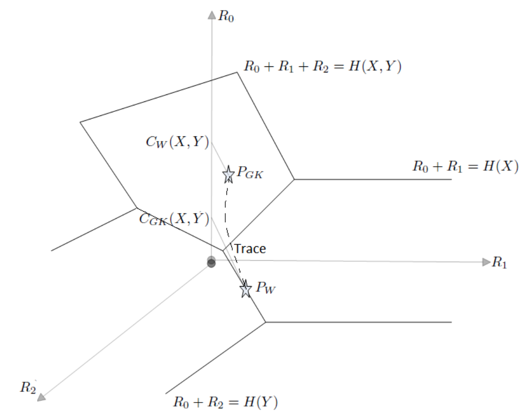


Figure 2: P_W and P_{GK} in the Gray-Wyner region. Observe that the transmit contour and the receive contour coincide in between P_W and P_{GW}.

where sup is taken over all f_1 and f_2 such that $P(f_1(X^n) \neq f_2(Y^n)) \to 0$ and where X^n and Y^n denote n independent copies of X and Y respectively. Let $\mathcal{X} \times \mathcal{Y} = \bigcup_{j=1}^{M} \mathcal{X}_j \times \mathcal{Y}_j$ be the ergodic decomposition of the stochastic matrix $P(X = x, Y = y)$. Define the random variable J as $J = j$ iff $x \in \mathcal{X}_j \Leftrightarrow y \in \mathcal{Y}_j$. Gács and Körner showed that $C_{GK}(X, Y) = H(J)$. Wyner also summarized the inequality relations between the quantities:

$$0 \leq C_{GK}(X, Y) \leq I(X; Y) \leq C_W(X, Y) \leq H(X, Y) \tag{15}$$

Gács and Körner's original definition of common information was naturally unrelated to the Gray-Wyner network, which appeared later. However, an alternate and insightful characterization of $C_{GK}(X, Y)$ was given by Ahlswede and Körner [4] in terms of $\mathcal{R}_{GW}$ as follow:

$$C_{GK}(X, Y) = \max R_0 : (R_0, R_1, R_2) \in \mathcal{R}_{GW} \tag{16}$$

subject to,

$$R_0 + R_1 = H(X), \;\; R_0 + R_2 = H(Y) \tag{17}$$

We denote the operating point in $\mathcal{R}_{GW}$ corresponding to $C_{GK}(X, Y)$, i.e. $(R_0, R_1, R_2) = (H(J), H(X) - H(J), H(Y) - H(J))$ by P_{GK}. It is again easy to observe that the quantity $K(X, Y; R)$ is a direct generalization of the Gács and Körner definition of common information to the '*excess receive rate*' regime. Clearly, we have $C_{GK}(X, Y) = K(X, Y; 0)$ as setting $R^{''} = 0$ in (5) forces both the constraint in (17).

It is well known that P_W and P_{GK} are two special points in $\mathcal{R}_{GW}$ as shown in Figure 2. However, no contour with practical significance is known which connects these points. The following claim sheds further light towards this understanding.

Claim 1. The transmit contour and the receive contour coincide with each other in between P_W and P_{GK}.

Proof: $\mathcal{R}_{GW}$ is a convex region. Hence the set of achievable rate pairs for $(R_t, R_r) = (R_0 + R_1 + R_2, 2R_0 + R_1 + R_2)$ is convex. We have $R_t \geq H(X, Y)$ and $R_r \geq H(X) + H(Y)$.

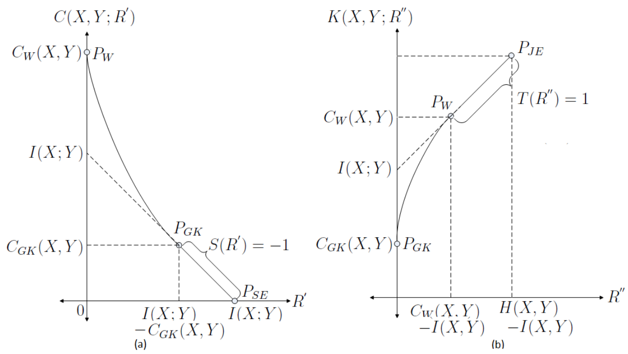


Figure 3: (a) Typical transmit tradeoff curve - $C(X,Y;R')$ (b) Typical receive tradeoff curve - $K(X,Y;R'')$

Note that when $R_t = H(X,Y)$, $\min R_r = H(X,Y) + C_W(X,Y)$ and is achieved at P_W. Similarly when $R_r = H(X) + H(Y)$, $\min R_t = H(X) + H(Y) - C_{GK}(X,Y)$, which is achieved at P_{GK}. Recall that the transmit contour is obtained by minimizing R_r for a fixed R_t and the receive contour is obtained by minimizing R_t at a fixed R_r. Hence, it follows from the convexity of (R_t, R_r) region that for every excess transmit rate $R' : 0 \leq R' \leq I(X;Y) - C_{GK}(X,Y)$, there exists an excess receive rate $R'' : 0 \leq R'' \leq C_W(X,Y) - I(X;Y)$ such that:

$$\begin{aligned} C(X,Y;R') &= K(X,Y;R'') \\ P_{C(X,Y)}(R') &= P_{K(X,Y)}(R'') \end{aligned} \qquad (18)$$

It follows that the transmit and the receive contours coincide in between P_W and P_{GK}. ■

This new relation between the two notions of common information brings both the definitions under a common framework. Similar to Wyner's definition, Gács-Körner's operating point can now be stated as the minimum shared rate at a sufficiently large sum rate. Likewise, Wyner's operating point can be defined as the maximum shared rate at a sufficiently large receive rate. We will make these arguments more precise in Section IV-A when we derive alternate characterizations for each notion in terms of the objective function of the other.

IV. Properties of the Tradeoff Curve

In this section, we shift our attention to the quantities $C(X,Y;R')$ and $K(X,Y;R'')$ instead of their contours and analyze some important properties. As most of the properties and their proofs for $K(X,Y;R'')$ are very similar to that for $C(X,Y;R')$, we only prove the properties for $C(X,Y;R')$. We plot typical transmit and receive tradeoff curves in Figure 3 to illustrate the discussion.

Consider the transmit tradeoff curve. At $R' = 0$ we get the Wyner's operating point where the minimum shared information is given by $C_W(X,Y)$. This point is denoted by P_W in Figure 3. Next observe that at $R_0 = 0$, for lossless reconstruction of X and Y, we need, $R_1 \geq H(X)$ and $R_2 \geq H(Y)$. Therefore at an excess sum rate $R' = H(X) + H(Y) - H(X,Y) = I(X;Y)$, the shared rate vanishes, or, $C(X,Y;I(X;Y)) = 0$. We call this point - 'separate encoding' and denote it by P_{SE} in the figure. It is also obvious that any U independent of (X,Y) achieves this minimum R_0 for $R' = I(X;Y)$.

Lemma 1. ***Convexity:*** *$C(X,Y;R')$ is convex $\forall R' \in [0, I(X;Y)]$ and $K(X,Y;R'')$ is concave $\forall R'' \in [0, H(X,Y) - I(X;Y)]$.*

Proof: The proof follows directly from the convexity of the Gray-Wyner region. ■

Lemma 2. ***Monotonicity*** *:$C(X,Y;R')$ is strictly monotone decreasing $\forall R' \in [0, I(X;Y)]$ and $K(X,Y;R'')$ is strictly monotone increasing $\forall R'' \in [0, H(X,Y) - I(X;Y)]$*

Proof: It is clear from the achievability results of Gray-Wyner that if a point $(r_0, r_1, r_2) \in \mathcal{R}_{GW}$, then all points $\{(R_0, R_1, R_2) : R_0 \geq r_0, R_1 \geq r_1, R_2 \geq r_2\} \in \mathcal{R}_{GW}$. Let us say, for some excess transmission rate R, $C(X,Y;R') = r_0$. Let the corresponding operating point in $\mathcal{R}_{GW}$ be (r_0, r_1, r_2). Hence for any $\Delta > 0$, the point $(r_0, r_1 + \Delta, r_2) \in \mathcal{R}_{GW}$ and satisfies $R_0 + R_1 + R_2 = R + \Delta$. Therefore,

$$\begin{aligned} C(X,Y;R' + \Delta) &= \min R_0 : \{R_0 + R_1 + R_2 = R + \Delta\} \\ &\leq r_0 \end{aligned} \qquad (19)$$

Hence, $C(X,Y;R')$ is non-increasing. Then it follows from convexity that $C(X,Y;R')$ is either a constant or is strictly monotone decreasing. Lemma 3 below eliminates the possibility of a constant, proving this lemma. ■

At all R' where $C(X,Y;R')$ is differentiable, we denote the slope by $S(R')$. At non-differentiable points, we denote by $S^-(R')$ and $S^+(R')$ the left and right derivatives respectively. Similarly the slope, left derivative and right derivatives of $K(X,Y;R'')$ are denoted by $T(R'')$, $T^-(R'')$ and $T^+(R'')$ respectively.

Lemma 3. *The slope of $C(X,Y;R')$, $S(R') \leq -1$ $\forall R \in [0, I(X;Y)]$ where the curve is differentiable. At non-differentiable points, we have $S^-(R') < S^+(R') \leq -1$. Similarly we have, $T(R'') \geq 1$ $\forall R'' \in [0, H(X,Y) - I(X;Y)]$ and $T^-(R'') > T^+(R'') \geq 1$*

Proof: Note that it is sufficient for us to show that $S^-(I(X;Y)) \leq -1$. Then it directly follows from convexity that $S(R') \leq -1$ at all differentiable points and $S^-(R') < S^+(R') \leq -1$ at all non-differentiable points. Consider $\triangle > 0$, and fix the shared information rate to be $R_0 = \triangle$. From the converse of the source coding theorem for lossless reconstruction, we have:

$$\begin{aligned} R_0 + R_1 = \triangle + R_1 &\geq H(X) \\ R_0 + R_2 = \triangle + R_2 &\geq H(Y) \end{aligned} \qquad (20)$$

The above inequalities imply $R_0 + R_1 + R_2 \geq H(X,Y) + I(X;Y) - \triangle$. Therefore the point on the transmit tradeoff curve with $C(X,Y;R') = \triangle$ has $R' \geq I(X;Y) - \triangle$. Hence $S^-(I(X;Y)) \leq -1$ proving the Lemma. ■

A. Alternate characterizations for $C_{GK}(X,Y)$ and $C_W(X,Y)$

In this section, we provide alternate characterizations for $C_{GK}(X,Y)$ and $C_W(X,Y)$ in terms of $C(X,Y;R')$ and

$C(X,Y;R^{''})$ respectively.

Theorem 2. *An alternate characterization for Gács-Körner's common information is:*

$$C_{GK}(X,Y) = \max_{R^{'}:S^{+}(R^{'})=-1} C(X,Y;R^{'}) \quad (21)$$

If there exists no $R^{'}$ for which $S^{+}(R^{'}) = -1$, then, $C_{GK}(X,Y) = 0$. Similarly, an alternate characterization for Wyner's common information is :

$$C_W(X,Y) = \min_{R^{''}:T^{+}(R^{''})=1} K(X,Y;R^{''}) \quad (22)$$

If there exists no $R^{''}$ for which $T^{+}(R) = 1$, then, $C_W(X,Y) = H(X,Y)$.

Note that $C_{GK}(X,Y)$ corresponds to that excess sum rate which demarcates the region of $C(X,Y;R^{'})$ with slope < -1 to the region with slope equal to -1 and $C_W(X,Y)$ corresponds to that excess receive rate which demarcates the region of $K(X,Y;R^{''})$ with slope > 1 to the region with slope equal to 1.

Proof: We first assume that there exists some $R^* \in [0, I(X;Y))$, for which $S^{+}(R^*) = -1$. We must prove that $C(X,Y;R^*) = C_{GK}(X,Y)$. We denote this point by P_{GK} in the figure. Let $\tilde{R}$ be such that $R^* \leq \tilde{R} < I(X;Y)$ and let $\tilde{U}$ be the random variable which achieves the minimum shared information rate at $\tilde{R}$ in Theorem 1. Then it follows from Lemmas 1 and 3 that $S^{+}(\tilde{R}) = -1$. Then the point in the GW region corresponding to $\tilde{U}$ satisfies the following two conditions:

$$\begin{aligned} R_0 &= I(X,Y) - \tilde{R} \\ R_0 + R_1 + R_2 &= H(X,Y) + \tilde{R} \end{aligned} \quad (23)$$

Adding the two equations, we have $2R_0 + R_1 + R_2 = H(X) + H(Y)$, which implies that $R_0 + R_1 = H(X)$ and $R_0 + R_2 = H(Y)$. Therefore, the point corresponding to $\tilde{U}$ satisfies Gács-Körner constraints (17). Hence, it follows that, any $\tilde{R}$ such that $S^{+}(\tilde{R}) = -1$ leads to an operating point in the GW region which satisfies Gács-Körner constraints.

Next, we need to show the converse. Consider any point in the GW region satisfying Gács-Körner constraints. It can be written as,

$$\begin{aligned} R_0 &= I(X;Y) - \tilde{R} \\ R_1 &= H(X) - (I(X;Y) - \tilde{R}) \\ R_2 &= H(Y) - (I(X;Y) - \tilde{R}) \end{aligned} \quad (24)$$

for some $C_{GK}(X,Y) \leq \tilde{R} \leq I(X;Y)$. On summing the three equations, we have $R_0 + R_1 + R_2 = H(X,Y) + \tilde{R}$. It then follows from the convexity of $C(X,Y;R^{'})$ that $S^{+}(\tilde{R}) = -1$. Therefore, we have,

$$\begin{aligned} C(X,Y;R^*) &= I(X;Y) - R^* \\ &= I(X;Y) - \min_{\tilde{R}:S^{+}(\tilde{R})=-1} \tilde{R} \\ &= \max R_0 : \ (R_0, R_1, R_2) \text{ satisfies (17)} \\ &= C_{GK}(X,Y) \end{aligned} \quad (25)$$

proving the first part of the theorem. However, if there exists no $R^{'} \in [0; I(X;Y)]$ for which $S^{+}(R^{'}) = -1$, it implies that $\forall R^{'} \in [0; I(X;Y)),\ C(X;Y;R^{'}) > I(X;Y)$. Therefore (17) is not satisfied with equality for any $R^{'} \in [0; I(X;Y))$. Hence $C_{GK}(X,Y) = 0$. ■

Note that the characterizations of $C_{GK}(X,Y)$ and $C_W(X,Y)$ in Theorem 2 are of fundamentally different nature from that in (12) and (16). These definitions not only provide further insight into the understanding of shared information, but also play a crucial role in finding the minimum communication cost for networks when the cost of transmission on each link is a non-linear function of the rate, for example, in capacity constrained networks. Just to illustrate, consider a network setting as shown in Fig. 1b, where the cost of transmission from the collector to the sinks is significantly higher than that from source to the collector. If the source-collector branch has a capacity constraint of R_C, then it follows directly from Theorem 2 that, $\forall R_C \geq H(X) + H(Y) - C_{GK}(X,Y)$, the minimum total cost is just a function of $C_{GK}(X,Y)$ and R_C.

V. Conclusion

Motivated by its applications in databases and routing in networks, the tradeoff between the the total transmit versus receive rate for the Gray-Wyner network was studied. Information theoretic characterizations for the tradeoff curves were established. Two well known notions of common information (Wyner's and Gács-Körner's) were shown to arise as extreme special cases of this broader framework. Using this relation, alternate characterizations under a common framework were derived for the two notions of common information.

References

[1] P. Gács and J. Körner, "Common information is far less than mutual information", Problems of Control and Information Theory, Vol. 2, pp. 149-162, 1973.

[2] A. Wyner, "The common information of two dependent random variables", IEEE Trans. on Information Theory, vol. 21, pp. 163 – 179, Mar 1975.

[3] R. Gray and A. Wyner, "Source coding for a simple network", Bell systems technical report, Dec 1974.

[4] R. Ahlswede and J. Körner, "On common information and related characteristics of correlated information sources", preprint presented at the 7th Prague conference on Information Theory, 1974.

[5] D. Marco and M. Effros "On lossless coding with coded side information", IEEE Trans. on Information Theory, 55 (7). pp. 3284-3296, Jul 2009

[6] S. Kamath and V. Anantharam, "A new dual to the Gács - Körner common information defined via the Gray-Wyner system," in Proc. 48th Allerton Conf. on Communication, Control, and Computing, pp. 1340–1346, 2010.

[7] V. M. Prabhakaran and M. M. Prabhakaran, "Assisted Common Information: Further Results." Available at: http://arxiv.org/abs/1105.6163

[8] H.Yamamoto, "Coding theorems for Shannon's cipher system with correlated source outputs, and common informations", IEEE Trans. on Infom. Theory, vol.40, no.1, pp.85-95, Jan. 1994.

[9] K. Viswanatha, E. Akyol and K. Rose, "On optimum communication cost for joint compression and dispersive information routing", IEEE Information theory workshop, pp. 1-5, Sep 2010.

[10] J. Nayak, S. Ramaswamy, and K. Rose, "Correlated source coding for fusion storage and selective retrieval," IEEE International Symposium on Information Theory, pp. 92-96, Sep 2005