ON THE CAPACITY REGION OF SOME CLASSES OF INTERFERENCE CHANNELS

Abstract. In this paper, we provide a new outer bound to the capacity region of the Gaussian Z-interference channel and also characterize the capacity of two new classes of discrete memoryless interference channels. The latter is achieved by proving the optimality of the Han-Kobayashi inner bound via traditional converse proofs. However, our outer bound for the Gaussian Z-interference channel is derived based on a novel observation. The conventional Gallager-type infeasibility proof identifies auxiliary random variables as the past/future of the channel random variables. Once an outer bound is established, it is important to establish a bound on the size of the alphabet of the auxiliary random variables to make the bound computable and meaningful. However, in the absence of cardinality bounds, we highlight that for Gaussian noise channels, if one can demonstrate the optimality of jointly Gaussian input distributions, such upper bounds become valuable. We demonstrate this point by utilizing an outer bound that is generally not computable for discrete interference channels to derive a new outer bound for Gaussian Z-interference channels.

1. Introduction

In this paper, we consider the two-sender-two-receiver interference channel model illustrated in Figure 1. Each sender wishes to transmit a message to its respective receiver via a shared medium. This channel models communication in two nearby cells in a wireless system. An (n, R_1, R_2) code $\mathcal{C}$, for the interference channel, consists of

- two message sets $[1 : 2^{nR_1}] := \{1, 2, ..., \lfloor 2^{nR_1} \rfloor\}$ and $[1 : 2^{nR_2}] := \{1, 2, ..., \lfloor 2^{nR_2} \rfloor\}$,
- two encoder functions $[1 : 2^{nR_i}] \to \mathcal{X}_i^n, i \in \{1, 2\}$ mapping each message m_i to a codeword x_i^n, and
- two decoder functions $\mathcal{Y}_i^n \to [1 : 2^{nR_i}], i \in \{1, 2\}$ mapping a codeword y_i^n to a message estimate, $\hat{m}_i$.

Assume that the messages (M_1, M_2) are uniformly distributed over $[1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$. The average probability error is defined to be

$$P_e^{(n)} = \Pr((\hat{M}_1, \hat{M}_2) \neq (M_1, M_2)).$$

In the traditional vanishing error setting, a rate pair (R_1, R_2) is *achievable* if there is a sequence of (n, R_1, R_2) codes such that $P_e^{(n)} \to 0$ as $n \to \infty$. Consider the region of achievable (R_1, R_2) rate-pairs. The closure of this achievable region is defined to be the capacity region $\mathscr{C}$. The reader can refer to [1, Chapter 6] for an overview of interference channels.

The capacity region is not known in general. The best known inner bound to the capacity region is the Han-Kobayashi inner bound.

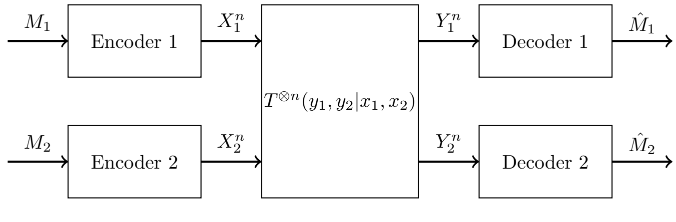

Figure 1. Interference Channel

Theorem 1 (Han-Kobayashi inner bound)**.** *Han-Kobayashi inner bound says that the following region is achievable for any* $p(q, u_1, u_2, x_1, x_2) = p(q)p(u_1, x_1|q)p(u_2, x_2|q)$*:*

$$\begin{aligned}
R_1 &< I(X_1; Y_1|U_2, Q),\\
R_2 &< I(X_2; Y_2|U_1, Q),\\
R_1 + R_2 &< I(X_1, U_2; Y_1|Q) + I(X_2; Y_2|U_1, U_2, Q),\\
R_1 + R_2 &< I(X_2, U_1; Y_2|Q) + I(X_1; Y_1|U_1, U_2, Q),\\
R_1 + R_2 &< I(X_1, U_2; Y_1|U_1, Q) + I(X_2, U_1; Y_2|U_2, Q),\\
2R_1 + R_2 &< I(X_1, U_2; Y_1|Q) + I(X_1; Y_1|U_1, U_2, Q) + I(X_2, U_1; Y_2|U_2, Q),\\
R_1 + 2R_2 &< I(X_2, U_1; Y_2|Q) + I(X_2; Y_2|U_1, U_2, Q) + I(X_1, U_2; Y_1|U_1, Q).
\end{aligned} \tag{1}$$

Moreover, the following cardinality bounds on the auxiliary random variables can be imposed: $|\mathcal{U}_1| \le |\mathcal{X}_1| + 4$, $|\mathcal{U}_2| \le |\mathcal{X}_2| + 4$ *and* $|\mathcal{Q}| \le 6$.

In general, the Han-Kobayashi inner bound is not equal to the capacity region. However, for certain special classes of interference channels, the Han-Kobayashi inner bound can indeed be optimal. One interesting case is the Gaussian Z-interference channels, where it remains an open question whether the Han-Kobayashi inner bound is tight. The current state-of-the-art outer bound does not match the Han-Kobayashi inner bound. In Section 2, we tighten the gap between the inner and outer bounds by providing a new outer bound. Our key observation is as follows: In the traditional Gallager-style infeasibility proof, it is common to identify auxiliary random variables as either the past or future (or both) of the channel random variables. After establishing an outer bound, it becomes essential to check if one can prove bounds on the size of the alphabet for the auxiliary random variables.[1] This step is necessary to ensure that the bound is computable. Without cardinality bounds, the outer bound region has no utility for discrete channels. To demonstrate the cardinality bounds on the auxiliary random variables, the Caratheodory-Bunt theorem is commonly employed. However, the effectiveness of this theorem is heavily contingent on the specific structure of the mutual information terms that manifest within the expression. The key idea of this paper is that outer bounds without cardinality bounds might be useful in Gaussian noise channels, even though they are not useful in discrete channels. This is because we possess alternative techniques that establish the optimality of jointly Gaussian random variables, which operate independently of the Caratheodory-Bunt theorem. Our example illustrates this point.

Finally, in Section 3 we show that the Han-Kobayashi inner bound is tight for two new classes of interference channels, including interference channels with strong interference at one receiver and injective deterministic function at the other.

2. Gaussian Z-interference

Consider the two-user Z-Gaussian interference channel:

$$\begin{aligned}
Y_1 &= X_1 + Z_1\\
Y_2 &= aX_1 + X_2 + Z_2\,,
\end{aligned} \tag{2}$$

with $a \in (0, 1)$, $Z_i \sim \mathcal{N}(0, 1)$ and a power constraint on the n-letter codebooks:

$$\|X_1^n\|^2 \le nP_1, \quad \|X_2^n\|^2 \le nP_2. \tag{3}$$

See Figure 2 for an illustration. If $a \ge 1$, the capacity region is fully known [2], [3] (strong interference regime). The capacity region is unknown when $0 < a < 1$ which corresponds to the weak interference regime.

An outer bound on the Z-interference channel was proven in [4]. The bound uses an auxiliary receiver J. We report the bound here for the choice of $J = Y_1$ as the simulation results in [4] indicate that the choice of $J = Y_1$ is optimal when we restrict to the Gaussian Z-interference channel.

[1] If computability were not a concern, one could write bounds with many auxiliary random variables, such as one for the past or future of each channel variable, and impose all possible relations between these auxiliary random variables that arise from rate constraints or the Csiszar sum lemma.

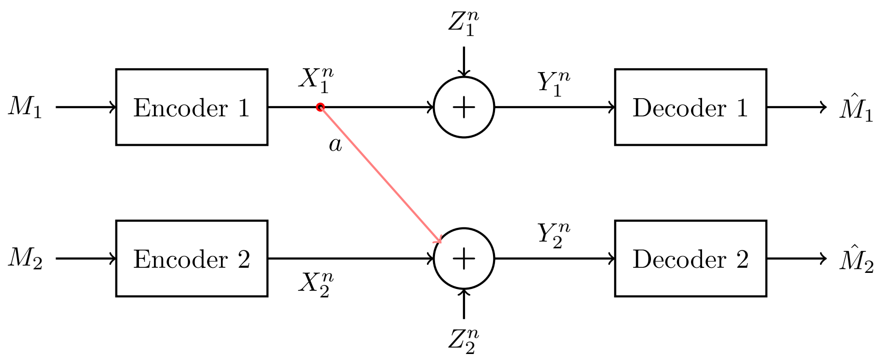

FIGURE 2. Illustration of a Gaussian Z-interference channel.

Theorem 2. *Consider the Z-interference channel* $p_{Y_1Y_2|X_1X_2} = p_{Y_1|X_1}p_{Y_2|Y_1,X_2}$. *Then, any rate pair* (R_1, R_2) *in the capacity of the Z-interference channel must satisfy the following constraints,*

$$R_1 \leq \min\{I(X_1;Y_1|Q), I(S;Y_2|Q) + I(X_1;Y_1|S,Q)\}$$
$$R_2 \leq \min\{I(X_2;Y_2|S,X_1,Q), I(X_2;Y_2|S,Q) - I(X_2;Y_1|S,Q)\}$$

for some auxiliary random variables S *and* Q *satisfying* $p(q)p(x_1|q)p(x_2|q)p(s|x_1,x_2,q)$ *and*

$$I(X_1;Y_1|S,Q) \geq I(X_1;Y_2|S,Q).$$

The proof uses the following choice for the auxiliary random variable:

$$S_i = (Y_2^{i-1}, Y_{1,i+1}^n).$$

In the following, we first state an outer bound for any arbitrary Z-interference channel. This upper bound is not computable since it provides no cardinality bound on its auxiliary random variables. However, we show later that for the Gaussian z-interference channel, the auxiliary random variables can be assumed to be jointly Gaussian and the outer bound becomes useful.

The following outer bound involves a receiver T. To understand the intuition for this receiver, observe that for the Gaussian interference channel given in (2), we can define $T = X_2 + Z_2$. Then, $Y_2 = aX_1 + T$ implies an injective deterministic situation, i.e. $T = g(Y_2, X_1), Y_2 = f(T, X_1)$.

Theorem 3. *Take an arbitrary Z-interference channel* $p(y_1|x_1)p(y_2|x_1,x_2)$ *such that*

$$I(A;Y_2|X_2) \leq I(A;Y_1|X_2)$$

for any $p(a,x_1,x_2)p(y_1,y_2|x_1,x_2)$, *i.e.,* Y_2 *is less noisy than* Y_1 *conditioned on* X_2.

Consider an auxiliary receiver T *such that* $p(y_1,y_2,t|x_1,x_2) = p(y_1|x_1)p(t|x_2)p(y_2|t,x_1)$ *and moreover* $H(T|Y_2,X_1) = H(Y_2|T,X_1) = 0$. *Then, for every* $\epsilon > 0$, *the following bound holds:*

$$R_1 \leq \epsilon + I(X_1;Y_1|U,Q)$$
$$R_2 \leq \epsilon + I(X_2;T|V,Q)$$

for some $p(q,u,v,w,x_1,x_2)$ *satisfying*

$$p(q,u,v,w,x_1,x_2) = p(v,x_2|q)p(u,x_1|q)p(w|v,u,x_1,q) \tag{4}$$

and the constraints:

$$I(W;Y_1|U,Q) = I(U;Y_2|W,Q)$$
$$I(V,W;Y_1|U,Q) = I(U;T,Y_2|V,W,Q)$$
$$I(X_1;Y_1|U,W,Q) \geq I(X_1;Y_2|U,W,Q).$$
$$I(X_2;T|V,Q) \leq I(X_2;Y_2|U,W,Q) - I(X_2;Y_1|U,W,Q) + \epsilon.$$

Remark 1. We cannot set $\epsilon = 0$ in the statement of the above bound (even for finite alphabet channels) because we cannot use the existing techniques to prove cardinality bounds on the alphabet of the auxiliary random variables.

Proof. Take a code of length n with input sequences $X_1^n = \mathcal{E}_1(M_1)$ and $X_2^n = \mathcal{E}_2(M_2)$. One can prove the above theorem by identifying the auxiliary variable

$$\begin{aligned} U_i &= Y_{1,i+1}^n, \\ V_i &= T^{i-1}, \\ W_i &= Y_2^{i-1}. \end{aligned}$$

Let $Q \in [1:n]$ be a time-sharing random variable, and let $U = (U_Q, Q)$, $V = (V_Q, Q)$, and $W = (W_Q, Q)$.

Observe that by Fano's inequality we have $n(R_1 - \epsilon_n) \leq I(M_1; Y_1^n) \leq I(X_1^n; Y_1^n)$ and $n(R_2 - \epsilon_n) \leq I(M_2; Y_2^n) \leq I(X_2^n; Y_2^n)$ for some ϵ_n that tends to zero as n tends to infinity. Now, observe that

$$I(X_1^n; Y_1^n) = \sum_i I(X_1^n; Y_{1i}|Y_1^{i-1}) = \sum_i I(X_{1i}; Y_{1i}|Y_1^{i-1})$$

and the first bound $R_1 \leq \epsilon + I(X_1; Y_1|U, Q)$ follows. The second inequality $R_2 \leq \epsilon + I(X_2; T|V, Q)$ follows from

$$\begin{aligned} I(X_2^n; Y_2^n) &\leq I(X_2^n; Y_2^n|X_1^n) \\ &= I(X_2^n; T^n|X_1^n) \\ &= I(X_2^n; T^n) \\ &= \sum_i I(X_{2i}; T_i|T^{i-1}) \\ &= nI(X_{2Q}; T_Q|V, Q). \end{aligned}$$

Next, we have

$$\begin{aligned} nI(X_{2Q}; T_Q|V, Q) &= I(X_2^n; Y_2^n|X_1^n) && (5) \\ &= I(X_2^n; Y_2^n) + I(X_2^n; X_1^n|Y_2^n) && (6) \\ &= I(X_2^n; Y_2^n) - I(X_2^n; Y_1^n) + I(X_2^n; X_1^n|Y_2^n) && (7) \\ &= \sum_i I(X_2^n; Y_{2i}|Y_{1,i+1}^n Y_2^{i-1}) - I(X_2^n; Y_{1i}|Y_{1,i+1}^n Y_2^{i-1}) + I(X_2^n; X_1^n|Y_2^n) && (8) \\ &= \sum_i I(X_{2i}; Y_{2i}|Y_{1,i+1}^n Y_2^{i-1}) - I(X_{2i}; Y_{1i}|Y_{1,i+1}^n Y_2^{i-1}) + I(X_2^n; X_1^n|Y_2^n) \\ &+ \sum_i I(X_2^n; Y_{2i}|X_{2i} Y_{1,i+1}^n Y_2^{i-1}) - I(X_2^n; Y_{1i}|X_{2i} Y_{1,i+1}^n Y_2^{i-1}) && (9) \\ &\leq \sum_i I(X_{2i}; Y_{2i}|Y_{1,i+1}^n Y_2^{i-1}) - I(X_{2i}; Y_{1i}|Y_{1,i+1}^n Y_2^{i-1}) + I(X_2^n; X_1^n|Y_2^n) && (10) \\ &\leq \sum_i I(X_{2i}; Y_{2i}|W_i, U_i) - I(X_{2i}; Y_{1i}|W_i, U_i) + n\epsilon_n && (11) \end{aligned}$$

where (10) follows from the less noisy property conditioned on X_2.

Since $p(y_1, t|x_1, x_2) = p(y_1|x_1)p(t|x_2)$, $I(X_1^n; X_2^n) = 0$ and $H(Y_2|T, X_1) = 0$, we obtain the Markov chain conditions in (4).

The constraints

$$\begin{aligned} I(W; Y_1|U, Q) &= I(U; Y_2|W, Q), \\ I(V, W; Y_1|U, Q) &= I(U; T, Y_2|V, W, Q) \end{aligned}$$

follow from the Csiszar-sum identity.

□

Next, we have the following result for Gaussian noise channels. Note that a natural way to characterize a convex region $\mathscr{R}$ is via its supporting hyperplanes, or equivalently via weighted-sum-rates. The weighted-sum-rate is defined as

$$\sup\{R_1 + \lambda R_2 : (R_1, R_2) \in \mathscr{R}\}$$

where $\lambda \geq 0$.

Theorem 4. *Take the Gaussian interference channel given in (2). Assume that $0 < a < 1$. Define $T = X_2 + Z_2$. Then, for any achievable rate pair (R_1, R_2) and $\lambda \geq 1$, we have*

$$R_1 + \lambda R_2 \leq I(X_1; Y_1|U) + \lambda I(X_2; T|V)$$

for some distribution p_{U,V,W,X_1,X_2} such that

$$p(u, v, w, x_1, x_2) = p(v, x_2)p(u, x_1)p(w|v, u, x_1) \tag{12}$$

and $p_{X_1|U}, p_{X_2|V}, p_{X_1,X_2|W}, p_{X_1,X_2|U,W}, p_{X_1,X_2|U,V,W}$ are all conditional Gaussian distributions whose variances (or covariances) do not depend on the conditioned variables (e.g., the variance of X_1 given $U = u$ does not depend on u); moreover, the following constraints are satisfied:

$$I(W; Y_1|U) = I(U; Y_2|W), \tag{13}$$

$$I(V, W; Y_1|U) = I(U; T, Y_2|V, W), \tag{14}$$

$$I(X_1; Y_1|U, W) \geq I(X_1; Y_2|U, W), \tag{15}$$

$$I(X_2; T|V) \leq I(X_2; Y_2|U, W) - I(X_2; Y_1|U, W), \tag{16}$$

hold. Further X_1 and X_2 are assumed to satisfy the power constraints, i.e. $\mathrm{E}\, X_1^2 \leq P_1$ and $\mathrm{E}\, X_2^2 \leq P_2$.

Remark 2. Let K_V, K_U, $K_{U,W}$ $K_{U,V,W}$ to be 2×2 the covariance matrix of X_1, X_2 given V, U, (U, W) and (U, V, W) respectively. Then, all the mutual information terms can be computed in terms of these covariance matrices, indicating that the outer bound is now in a computable and compact form. To evaluate the bound, one has to take the union over all feasible covariance matrices. Besides the constraints like $K_{U,W} \preceq K_U$ coming from Jensen's inequality, the Markov structure imposes further some constraints on K_V, K_U, $K_{U,W}$ $K_{U,V,W}$. For instance, K_V, K_U are diagonal matrices. We also know that $\mathsf{var}(X_2|V) = \mathsf{var}(X_2|U, W, V, X_1) \leq \mathsf{var}(X_2|U, W, X_1)$.

Remark 3. The above theorem only considers $\lambda \geq 1$ because for $\lambda = 1$, the sum rate is known to be optimized at Sato's corner point, [2, 5]. Further, for any $\lambda \in (0, 1]$, the supporting hyperplane to the capacity region also passes through Sato's corner point.

Proof. We utilize the ideas in [6] to deduce the Gaussian extremality. First, let us consider the optimization problem induced by Theorem 3. That is, we want to maximize $I(X_1; Y_1|U, Q) + \lambda I(X_2; T|V, Q)$ over distributions of the form

$$p(q, u, v, w, x_1, x_2) = p(v, x_2|q)p(u, x_1|q)p(w|v, u, x_1, q), \tag{17}$$

satisfying the power constraints, i.e. $\mathrm{E}\, X_1^2 \leq P_1$ and $\mathrm{E}\, X_2^2 \leq P_2$ and satisfying the following constraints:

$$I(W; Y_1|U, Q) = I(U; Y_2|W, Q)$$

$$I(V, W; Y_1|U, Q) = I(U; T, Y_2|V, W, Q)$$

$$I(X_1; Y_1|U, W, Q) \geq I(X_1; Y_2|U, W, Q).$$

$$I(X_2; T|V, Q) \leq I(X_2; Y_2|U, W, Q) - I(X_2; Y_1|U, W, Q) + \epsilon.$$

It suffices to show that conditional Gaussian distributions maximize the outer bound in Theorem 3 for every $\epsilon > 0$. We can then let ϵ converge to zero and get the bound stated in the theorem. The random variable Q can be dropped at the very end as $I(X_1; Y_1|U, Q = q)$ will not depend on the value of q. Therefore, we can just consider the conditional distribution of all random variables given some $Q = q$.

We utilize the ideas in [6]. Gaussian optimality using these ideas requires the following components: the existence of a maximizer, a sub-additive function (that is additive for independent distributions and invariant under rotations), and equality in the sub-additivity holds only if the distributions are independent.

To show the existence of a maximizer, we use Prokhorov's theorem. We refer the reader to the arguments in Appendix II of [6] as the arguments are standard. To use Prokhorov's theorem, one needs to show the tightness of the distributions (for subsequential convergence) and continuity of differential entropy. To justify tightness in our problem, we note that without loss of generality, we can assume that random variables Q, U, V, W all take values in the compact set $[0, 1]$ (to see this, note that we can replace U with $f(U)$ where $f(\cdot)$ is the one-to-one mapping $f(x) = \frac{1}{1+e^x}$. None of the mutual information terms will be affected by this transformation). We also have power constraints on X_1, X_2. This will yield the requisite *tightness* condition on a sequence of distributions using Prokhorov's theorem. The additive Gaussian noise yields the requisite continuity of differential entropy. Together, we get the existence of maximizing distributions.

The key component is the second part, i.e., identifying the sub-additive function. Intuitively speaking, the outer bound derivations already provide such a sub-additivity argument. We provide the details here:

Take $(X_{11}^*, X_{21}^*, Q_1^*, U_1^*, V_1^*, W_1^*)$ and $(X_{12}^*, X_{22}^*, Q_2^*, U_2^*, V_2^*, W_2^*)$ to be two independent copies of a maximizer $p^*(x_1, x_2, q, u, v, w)$ of the optimization problem of Theorem 3 with the objective function

$$I(X_1; Y_1|U, Q) + \lambda I(X_2; T|V, Q),$$

subject to the constraints in the theorem. Further, let V_λ denote the optimal value. Let

$$(\hat{Q}, \hat{U}, \hat{V}, \hat{W}) := ((Q_1^*, Q_2^*), (U_1^*, U_2^*), (V_1^*, V_2^*), (W_1^*, W_2^*))$$

and

$$\begin{pmatrix} X_{1+} \\ X_{1-} \end{pmatrix} := \begin{pmatrix} \sqrt{t} & \sqrt{1-t} \\ -\sqrt{1-t} & \sqrt{t} \end{pmatrix} \begin{pmatrix} X_{11}^* \\ X_{12}^* \end{pmatrix}, \qquad \begin{pmatrix} X_{2+} \\ X_{2-} \end{pmatrix} := \begin{pmatrix} \sqrt{t} & \sqrt{1-t} \\ -\sqrt{1-t} & \sqrt{t} \end{pmatrix} \begin{pmatrix} X_{21}^* \\ X_{22}^* \end{pmatrix}.$$

where $t = \frac{1}{2}$. We define channel outputs $Y_{1+}, Y_{1-}, T_+, T_-, Y_{2+}, Y_{2-}$ in the standard manner as in [6].

Set

$$\hat{U}_- := (\hat{U}, Y_{1+}) \qquad \hat{U}_+ := \hat{U}, \tag{18}$$
$$\hat{V}_- := \hat{V} \qquad \hat{V}_+ := (\hat{V}, T_-) \tag{19}$$
$$\hat{W}_- := \hat{W} \qquad \hat{W}_+ := (\hat{W}, Y_{2-}) \tag{20}$$

Since rotations preserve mutual information, note that standard manipulations yield

$$\begin{aligned} 2V &= I(X_{1+}X_{1-}; Y_{1+}Y_{1-}|\hat{U}, \hat{Q}) + \lambda I(X_{2+}X_{2-}; T_+T_-|\hat{V}, \hat{Q}) \\ &= I(X_{1+}; Y_{1+}|\hat{U}, \hat{Q}) + \lambda I(X_{2+}; T_+|\hat{V}, T_-, \hat{Q}) \\ &\quad + I(X_{1-}; Y_{1-}|\hat{U}, Y_{1+}\hat{Q}) + \lambda I(X_{2-}; T_-|\hat{V}, \hat{Q}). \end{aligned}$$

Now consider a joint distribution of the form $p^*(s, q, u, v, w, x_1, x_2)$ where S takes values in $\{+, -\}$ with uniform probability. Conditioned on $S = +$, the rest of random variables are distributed according to $P_{\hat{Q}, \hat{U}_+, \hat{V}_+, \hat{W}_+, X_{1+}, X_{2+}}$; similarly, conditioned on $S = -$, the rest are distributed according to $P_{\hat{Q}, \hat{U}_-, \hat{V}_-, \hat{W}_-, X_{1-}, X_{2-}}$.

The following equations (or inequations) can be directly verified, showing that p^*_{S,Q,U,V,W,X_1,X_2} is another maximizer of Theorem 3:

$$\begin{aligned} I(X_{1+}X_{1-}; Y_{1+}Y_{1-}|\hat{U}, \hat{Q}) &= I(X_{1+}; Y_{1+}|\hat{U}_+, \hat{Q}) + I(X_{1-}; Y_{1-}|\hat{U}_-, \hat{Q}), \\ I(X_{2+}X_{2-}; T_+T_-|\hat{V}, \hat{Q}) &= I(X_{2+}; T_+|\hat{V}_+, \hat{Q}) + I(X_{2-}; T_-|\hat{V}_-, \hat{Q}), \\ I(\hat{W}; Y_{1+}Y_{1-}|\hat{U}, \hat{Q}) - I(\hat{U}; Y_{2+}Y_{2-}|\hat{W}, \hat{Q}) &= I(\hat{W}_+; Y_{1+}|\hat{U}_+, \hat{Q}) - I(\hat{U}_+; Y_{2+}|\hat{W}_+, \hat{Q}) \\ &\quad + I(\hat{W}_-; Y_{1-}|\hat{U}_-, \hat{Q}) - I(\hat{U}_-; Y_{2-}|\hat{W}_-, \hat{Q}), \\ I(\hat{V}, \hat{W}; Y_{1+}, Y_{1-}|\hat{U}, \hat{Q}) - I(\hat{U}; T_+, Y_{2+}, T_-, Y_{2-}|\hat{V}, \hat{W}, \hat{Q}) &= I(\hat{V}_+, \hat{W}_+; Y_{1+}|\hat{U}_+, \hat{Q}) - I(\hat{U}_+; T_+, Y_{2+}|\hat{V}_+, \hat{W}_+, \hat{Q}) \\ &\quad + I(\hat{V}_-, \hat{W}_-; Y_{1-}|\hat{U}_-, \hat{Q}) - I(\hat{U}_-; T_-, Y_{2-}|\hat{V}_-, \hat{W}_-, \hat{Q}). \end{aligned}$$

Next, we have

$$\begin{aligned}
0 &\le I(X_{1+}X_{1-};Y_{1+}Y_{1-}|\hat{U},\hat{W},\hat{Q}) - I(X_{1+}X_{1-};Y_{2+}Y_{2-}|\hat{U},\hat{W},\hat{Q}) \\
&= I(X_{1+}X_{1-};Y_{1+}|\hat{U}_+,\hat{W}_+,\hat{Q}) - I(X_{1+}X_{1-};Y_{2+}|\hat{U}_+,\hat{W}_+,\hat{Q}) \\
&\quad + I(X_{1+}X_{1-};Y_{1-}|\hat{U}_-,\hat{W}_-,\hat{Q}) - I(X_{1+}X_{1-};Y_{2-}|\hat{U}_-,\hat{W}_-,\hat{Q}) \\
&= I(X_{1+};Y_{1+}|\hat{U}_+,\hat{W}_+,\hat{Q}) - I(X_{1+}X_{1-};Y_{2+}|\hat{U}_+,\hat{W}_+,\hat{Q}) \qquad (21) \\
&\quad + I(X_{1-};Y_{1-}|\hat{U}_-,\hat{W}_-,\hat{Q}) - I(X_{1+}X_{1-};Y_{2-}|\hat{U}_-,\hat{W}_-,\hat{Q}) \qquad (22) \\
&\le I(X_{1+};Y_{1+}|\hat{U}_+,\hat{W}_+,\hat{Q}) - I(X_{1+};Y_{2+}|\hat{U}_+,\hat{W}_+,\hat{Q}) \\
&\quad + I(X_{1-};Y_{1-}|\hat{U}_-,\hat{W}_-,\hat{Q}) - I(X_{1-};Y_{2-}|\hat{U}_-,\hat{W}_-,\hat{Q}).
\end{aligned}$$

where in (21) and (22) we use the fact that Y_1 depends only on X_1.

We have the constraint

$$I(X_{2+}X_{2-};T_+T_-|\hat{V},\hat{Q}) \le I(X_{2+}X_{2-};Y_{2+}Y_{2-}|\hat{U},\hat{W},\hat{Q}) - I(X_{2+}X_{2-};Y_{1+}Y_{1-}|\hat{U},\hat{W},\hat{Q}).$$

On the one hand,

$$I(X_{2+}X_{2-};T_+T_-|\hat{V},\hat{Q}) = I(X_{2+};T_+|\hat{V}_+,\hat{Q}) + I(X_{2-};T_-|\hat{V}_-,\hat{Q})$$

and on the other hand we have

$$\begin{aligned}
&I(X_{2+}X_{2-};Y_{2+}Y_{2-}|\hat{U},\hat{W},\hat{Q}) - I(X_{2+}X_{2-};Y_{1+}Y_{1-}|\hat{U},\hat{W},\hat{Q}) \\
&= I(X_{2+}X_{2-};Y_{2+}|\hat{U}_+,\hat{W}_+,\hat{Q}) - I(X_{2+}X_{2-};Y_{1+}|\hat{U}_+,\hat{W}_+,\hat{Q}) \\
&\quad + I(X_{2+}X_{2-};Y_{2-}|\hat{U}_-,\hat{W}_-,\hat{Q}) - I(X_{2+}X_{2-};Y_{1-}|\hat{U}_-,\hat{W}_-,\hat{Q}) \\
&= I(X_{2+};Y_{2+}|\hat{U}_+,\hat{W}_+,\hat{Q}) - I(X_{2+};Y_{1+}|\hat{U}_+,\hat{W}_+,\hat{Q}) \\
&\quad + I(X_{2-};Y_{2-}|\hat{U}_-,\hat{W}_-,\hat{Q}) - I(X_{2-};Y_{1-}|\hat{U}_-,\hat{W}_-,\hat{Q}) \\
&\quad + I(X_{2-};Y_{2+}|X_{2+},\hat{U}_+,\hat{W}_+,\hat{Q}) - I(X_{2-};Y_{1+}|X_{2+},\hat{U}_+,\hat{W}_+,\hat{Q}) \\
&\quad + I(X_{2+};Y_{2-}|X_{2-},\hat{U}_-,\hat{W}_-,\hat{Q}) - I(X_{2+};Y_{1-}|X_{2-},\hat{U}_-,\hat{W}_-,\hat{Q}) \\
&\le I(X_{2+};Y_{2+}|\hat{U}_+,\hat{W}_+,\hat{Q}) - I(X_{2+};Y_{1+}|\hat{U}_+,\hat{W}_+,\hat{Q}) \qquad (23) \\
&\quad + I(X_{2-};Y_{2-}|\hat{U}_-,\hat{W}_-,\hat{Q}) - I(X_{2-};Y_{1-}|\hat{U}_-,\hat{W}_-,\hat{Q}) \qquad (24)
\end{aligned}$$

where (23) and (24) follow from the less noisy property conditioned on X_2.

Intuitively, one can see the Gaussian optimality as X_{1+}, X_{1-} edge closer to Gaussian by the Central limit theorem. However, it turns out to be easier to invoke the Darmois-Skitovic theorem as shown in [6].

Theorem 5 (Darmois-Skitovich theorem [7,8])**.** *Let $X_1,\dots,X_n$ be independent random variables. Further, let $\alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_n$ be non-zero constants for each coordinate. If the linear statistics $L_1 = \sum_{i=1}^n \alpha_i X_i$ and $L_2 = \sum_{i=1}^n \beta_i X_i$ are independent, then all random variables $X_1,\dots,X_n$ are Gaussians.*

Therefore, the third component of the technique in [6] is to show the independence of the rotated forms. Sometimes, one gets this directly via the proof of sub-additivity as in [6], or at other times one has to modify the sub-additive functional by a small amount to induce this independence (as illustrated in [4]). Here, we use the latter approach and add some small perturbation terms to our objective function $I(X_1;Y_1|U,Q)+\lambda I(X_2;T|V,Q)$.

The choice of the identification of the auxiliaries guide the design of the perturbative terms. It suffices to show that at the maximizer $p_{X_1|U}, p_{X_2|V}, p_{X_1,X_2|W}, p_{X_1,X_2|U,W}, p_{X_1,X_2|U,V,W}$ are all Gaussians whose variances (or covariances) do not depend on the conditioned variables. To show each of the following, we perturb our subadditive function appropriately using functions uniformly bounded by say δ, and conclude the Gaussians optimize these perturbed functions. Then, we let $\delta \to 0$ to deduce the Gaussian optimality for the original function.

Below, we outline the concept behind the construction of the perturbation terms. Consider the term $I(X_1;Y_1^\dagger|U)$, where $Y_1^\dagger$ is a degraded version (by the addition of Gaussian noise) of Y_1. Now observe that subadditivity for this term proceeds as follows:

$$\begin{aligned} I(X_1,X_2;Y_{11}^\dagger,Y_{12}^\dagger|U_1,U_2) &= I(X_{1+},X_{1-};Y_{1+}^\dagger,Y_{1-}^\dagger|\hat{U}) \\ &= I(X_{1+};Y_{1+}^\dagger|\hat{U}) + I(X_{1-};Y_{1-}^\dagger|\hat{U},Y_{1+}^\dagger) \\ &= I(X_{1+};Y_{1+}^\dagger|\hat{U}) + I(X_{1-};Y_{1-}^\dagger|\hat{U},Y_{1+}) + I(Y_{1+};Y_{1-}^\dagger|\hat{U},Y_{1+}^\dagger) \\ &= I(X_{1+};Y_{1+}^\dagger|\hat{U}_+) + I(X_{1-};Y_{1-}^\dagger|\hat{U}_-) + I(Y_{1+};Y_{1-}^\dagger|\hat{U},Y_{1+}^\dagger) \end{aligned}$$

Assume that we add $-\delta_1 I(X_1;Y_1^\dagger|U)$ to our original function

$$I(X_1;Y_1|U,Q) + \lambda I(X_2;T|V,Q) - \delta_1 I(X_1;Y_1^\dagger|U),$$

If we take the product of two maximizers, then the rotated versions will also be a maximizer and will strictly improve unless $I(Y_{1+};Y_{1-}^\dagger|\hat{U},Y_{1+}^\dagger)=0$. Note, by the degraded structure, we also have $I(Y_{1+}^\dagger;Y_{1-}^\dagger|\hat{U},Y_{1+})=0$, implying double Markovity and hence (see [4] for details of such arguments) Y_{1-} is independent of $(Y_{1+}^\dagger,Y_{1+})$, conditioned on U, and hence that X_{1+},X_{1-} are also independent conditioned on U. Then by applying Theorem 5, we have that, at the maximizer, $p_{X_1|U}$ is a Gaussian. Further, we also know that the variance does not depend on U (see Corollary 3 in [6].)

By considering the term $I(X_2;T^\dagger|V)$, where $T^\dagger$ is a degraded (by independent Gaussian noise) version of T_2, and doing the same operations as above, we can deduce that $p_{X_2|V}$ is Gaussian with a covariance that does not depend on V. The excess term we obtain from the subadditivity proof is $I(T_-;T_{2+}^\dagger|\hat{V},T_-)$ in this case.

Simiarly, by considering the term $I(X_1,X_2;Y_2^\dagger|W)$, where $Y_2^\dagger$ is a degraded (by independent Gaussian noise) version of Y_2, and doing the same operations as above, we can deduce that $p_{X_1,X_2|W}$ is Gaussian with a covariance that does not depend on W. The excess term we obtain here is $I(Y_{2-};Y_{2+}^\dagger|\hat{W},Y_{2-})$

Now consider the term $I(X_1,X_2;Y_1|U,W) - \mu I(X_1,X_2;Y_1|U,W)$, for some fixed μ satisfying $\mu > 1$. Performing similar manipulations as before, we obtain an excess term $(\mu-1)I(Y_{1+};Y_{2-}|U,W)$. This being zero at maximizer forces $p_{X_1,X_2|U,W}$ to be Gaussian with a covariance that does not depend on U,W.

Finally, consider the term $I(X_1,X_2;T,Y_2|U,W,V) - \mu I(X_1,X_2;Y_1|U,W,V)$, for some fixed μ satisfying $\mu > 1$. Again, performing similar manipulations as before, we obtain an excess term $(\mu-1)I(Y_{1+};T_-,Y_{2-}|U,V,W)$. This being zero at maximizer forces $p_{X_1,X_2|U,V,W}$ to be Gaussian with a covariance that does not depend on U,V,W.

□

2.1. **Comparison of the outer bounds.** We show that the bound in Theorem 4 is less than or equal to the bound in Theorem 2. To show this, it suffices to show that for any $\lambda \geq 1$, the bound on the λ-sum rate in Theorem 4 is less than or equal to that in Theorem 2.

Take a point in the bound in Theorem 4. Corresponding to the point are auxiliary random variables U,V,W yielding the upper bound on the λ-sum rate. We show that $R_1 = I(X_1;Y_1|U)$ and $R_2 = I(X_2;T|V)$ belongs to the region in Theorem 2 and can be obtained via the auxiliary random variable $S=(U,W)$ and constant Q.

First, consider the bounds on R_1 in Theorem 2: it is immediate that

$$I(X_1;Y_1|U) \leq I(X_1;Y_1).$$

We also claim that

$$I(X_1;Y_1|U) \leq I(U,W;Y_2) + I(X_1;Y_1|U,W)$$

which shows that R_1 will be satisfied for $S=(U,W)$. Using the fact that $I(W;Y_1|U) = I(U;Y_2|W)$, we can write

$$\begin{aligned} I(U,W;Y_2) + I(X_1;Y_1|U,W) &= I(W;Y_2) + I(W;Y_1|U) + I(X_1;Y_1|U,W) \\ &= I(W;Y_2) + I(W,X_1;Y_1|U) \\ &\geq I(X_1;Y_1|U). \end{aligned} \tag{25}$$

Next, consider the bounds on R_2: since

$$I(X_2;T|V) \leq I(X_2;Y_2|U,W) - I(X_2;Y_1|U,W),$$

the second inequality on R_2 in Theorem 2 is satisfied with the auxiliary random variable $S = (U, W)$ and constant Q. It remains to show that

$$I(X_2; T|V) \leq I(X_2; Y_2|U, W, X_1).$$

Since $(X_2, T) \to V \to (U, W, X_1)$ forms a Markov chain, we can write

$$\begin{aligned} I(X_2; T|V) &= I(X_2; T|V, U, W, X_1) && (26)\\ &= I(X_2; Y_2|V, U, W, X_1) && (27)\\ &\leq I(V, X_2; Y_2|U, W, X_1) && (28)\\ &= I(X_2; Y_2|U, W, X_1). && (29)\end{aligned}$$

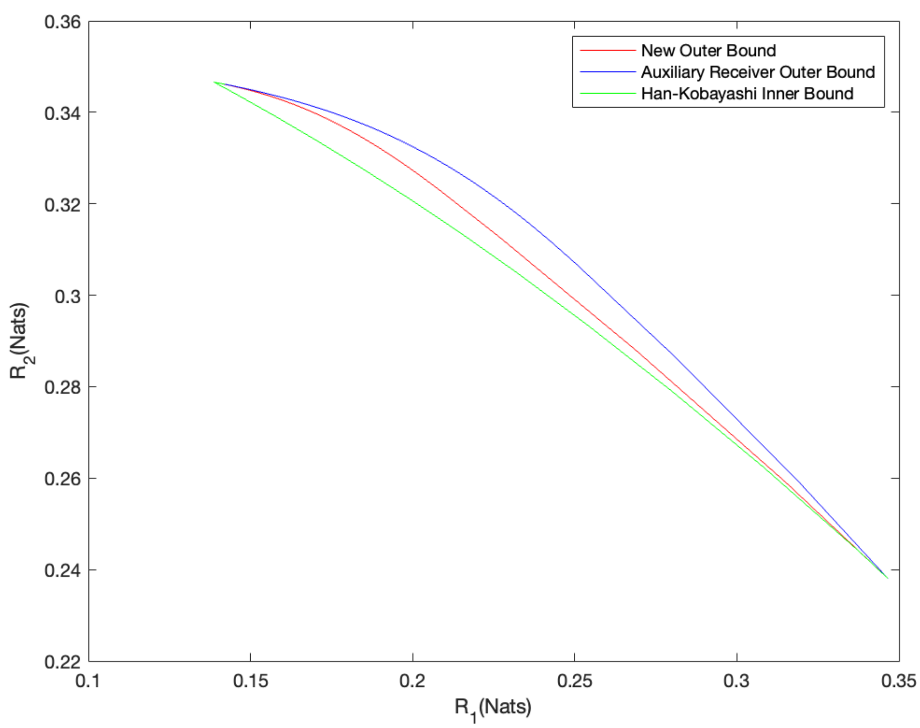

FIGURE 3. Comparison of the bounds when $a = 0.8, P_1 = 1, P_2 = 1$. The green curve is the Han-Kobayashi inner bound. The red curve is a relaxed version of the bound in Theorem 4, and the blue curve is the one obtained by Theorem 2.

3. THE CAPACITY OF TWO CLASSES OF DISCRETE INTERFERENCE CHANNELS

3.1. A class of interference channels with strong interference at one receiver and injective deterministic function at the other.

Definition 1. We say that the interference channel $p(y_1, y_2|x_1, x_2)$ has strong interference from the first transmitter if $I(X_1; Y_2|X_2) \geq I(X_1; Y_1|X_2)$ for every $p(x_1)p(x_2)$.

The following property is established in [1, p. 137]: if $p(y_1, y_2|x_1, x_2)$ has strong interference from the first transmitter, then

$$I(X_1^n; Y_2^n|X_2^n) \geq I(X_1^n; Y_1^n|X_2^n) \quad (30)$$

for every $p(x_1^n)p(x_2^n)$ (i.e., the condition tensorizes). Moreover, in any (n, ϵ_n) code, we have

$$n(R_1 + R_2) \leq \sum_{i=1}^{n} I(X_{1i}, X_{2i}; Y_{2i}) + n\epsilon_n. \tag{31}$$

Theorem 6. *Consider an interference channel where $Y_1 = g(T, X_1)$ for some function g where $T = f(X_2)$ is a function of X_2. Moreover, assume that $T = k(Y_1, X_1)$ for some function $k(\cdot, \cdot)$. One is also given that the interference channel has strong interference from the first transmitter (Definition 1). Then, the capacity region of this channel is as follows: a rate pair (R_1, R_2) is achievable if and only if it*

$$R_1 \leq I(X_1; Y_1|T, Q), \tag{32}$$
$$R_2 \leq I(X_2; Y_2|X_1, Q), \tag{33}$$
$$R_1 + R_2 \leq I(X_1, T; Y_1|Q) + I(X_2; Y_2|X_1, T, Q), \tag{34}$$
$$R_1 + R_2 \leq I(X_1, X_2; Y_2|Q) \tag{35}$$
$$R_1 + R_2 \leq I(T; Y_1|X_1, Q) + I(X_1, X_2; Y_2|T, Q), \tag{36}$$
$$2R_1 + R_2 \leq I(X_1, T; Y_1|Q) + I(X_1, X_2; Y_2|T, Q), \tag{37}$$

where $p(x_1, x_2, q) = p(q)p(x_1|q)p(x_2|q)$ for some $Q : |\mathcal{Q}| \leq 6$.

Proof. To show the achievability of the region, set $U_1 = X_1$ and $U_2 = T$ in the Han-Kobayashi inner bound in Theorem 1. This yields the region given in the theorem's statement except that the Han-Kobayashi inner bound has one extra inequality:

$$R_1 + 2R_2 \leq I(X_1, X_2; Y_2|Q) + I(X_2; Y_2|X_1, T, Q) + I(T; Y_1|X_1, Q). \tag{38}$$

However, the inequality (38) is redundant and implied by adding (33) and (35) as

$$I(X_2; Y_2|X_1, T, Q) + I(T; Y_1|X_1, Q) \geq I(X_2; Y_2|X_1, Q).$$

The above inequality holds since T is a function of (X_1, Y_1).

It remains to prove the converse. Note that the cardinality bound on Q comes from the standard Caratheodory-Bunt arguments, so we need to show that any achievable rate pair satisfies the inequalities for some $p(q)p(x_1|q)p(x_2|q)$.

To get a matching outer bound, we proceed as follows: for (32) we write

$$n(R_1 - \epsilon_n) \leq I(X_1^n; Y_1^n|T^n) = H(Y_1^n|T^n) = \sum_i H(Y_{1i}|Y_1^{i-1}T^n) \leq \sum_i H(Y_{1i}|T_i) = \sum_i I(X_{1i}; Y_{1i}|T_i).$$

Next, for (33) we write

$$\begin{aligned} n(R_2 - \epsilon_n) &\leq I(X_2^n; Y_2^n|X_1^n) = H(Y_2^n|X_1^n) - H(Y_2^n|X_1^n, X_2^n) = \sum_i H(Y_{2i}|Y_2^{i-1}X_1^n) - H(Y_{2i}|X_{1i}X_{2i}) \\ &\leq \sum_i H(Y_{2i}|X_{1i}) - H(Y_{2i}|X_{1i}X_{2i}) = \sum_i I(X_{2i}; Y_{2i}|X_{1i}). \end{aligned}$$

Next, for (34) we write

$$\begin{aligned} n(R_1 + R_2 - \epsilon_n) &\leq I(X_1^n; Y_1^n) + I(X_2^n; Y_2^n|X_1^n) + H(T^n|X_1^n) - I(T^n; Y_2^n|X_1^n) \\ &= I(X_1^n; Y_1^n) + I(T^n, X_2^n; Y_2^n|X_1^n) + I(T^n; Y_1^n|X_1^n) - I(T^n; Y_2^n|X_1^n) && (39) \\ &= I(X_1^n, T^n; Y_1^n) + I(X_2^n; Y_2^n|X_1^n, T^n) \\ &= H(Y_1^n) + H(Y_2^n|X_1^n, T^n) - \sum_i H(Y_{2i}|X_{1i}X_{2i}T_i) && (40) \\ &\leq \sum_i H(Y_{1i}) + H(Y_{2i}|X_{1i}, T_i) - H(Y_{2i}|X_{1i}X_{2i}T_i) && (41) \\ &= \sum_i I(X_{1i}, T_i; Y_{1i}) + I(X_{2i}; Y_{2i}|X_{1i}, T_i) && (42) \end{aligned}$$

where (39) follows from $H(T^n|X_2^n) = H(T^n|X_1^n, Y_1^n) = 0$ and (40) and (42) follow from Y_1 being a function of (T, X_1).

The inequality (35) follows from (31). For the inequality (36) we write

$$\begin{aligned} n(R_1+R_2-\epsilon_n) &\le I(X_2^n;Y_2^n)+I(X_1^n;Y_1^n|X_2^n) \\ &\le H(T^n)+I(X_2^n;Y_2^n|T^n)+I(X_1^n;Y_1^n|X_2^n) \\ &\le H(T^n)+I(X_2^n;Y_2^n|T^n)+I(X_1^n;Y_2^n|X_2^n) \quad (43) \\ &= H(T^n)+I(X_1^n,X_2^n;Y_2^n|T^n) \\ &= H(T^n)+H(Y_2^n|T^n)-\sum_i H(Y_{2i}|X_{1i},X_{2i},T_i) \\ &\le \sum_i H(T_i)+H(Y_{2i}|T_i)-\sum_i H(Y_{2i}|X_{1i},X_{2i},T_i) \\ &= \sum_i I(T_i;Y_{1i}|X_{1i})+I(X_{1i},X_{2i};Y_{2i}|T_i) \quad (44) \end{aligned}$$

where (43) follows from (30) and (44) follows from $H(T|Y_1,X_1)=H(T|X_2)=0$, implying that T_i is independent of X_{1i} for all i. Finally for (37), observe that

$$\begin{aligned} n(2R_1+R_2-\epsilon_n) &\le I(X_1^n;Y_1^n)+I(X_1^n;Y_1^n|X_2^n)+I(X_2^n;Y_2^n) \\ &\overset{(a)}{\le} I(X_1^n;Y_1^n)+I(X_1^n;Y_2^n|X_2^n)+I(X_2^n;Y_2^n) \\ &= I(X_1^n;Y_1^n)+I(X_1^n,X_2^n;Y_2^n) \\ &\overset{(b)}{=} I(X_1^n;Y_1^n)+I(X_1^n,X_2^n;Y_2^n|T^n)+I(T^n;Y_2^n) \\ &\overset{(c)}{\le} I(X_1^n;Y_1^n)+I(X_1^n,X_2^n;Y_2^n|T^n)+I(T^n;Y_1^n|X_1^n) \\ &= I(X_1^n,T^n;Y_1^n)+I(X_1^n,X_2^n;Y_2^n|T^n) \\ &\overset{(d)}{\le} \sum_i I(X_{1i},T_i;Y_{1i})+I(X_{1i},X_{2i};Y_{2i}|T_i). \end{aligned}$$

Here, (a) follows from the strong interference condition $I(X_1^n;Y_1^n|X_2^n)\le I(X_1^n;Y_2^n|X_2^n)$ (see (30)), (b) follows from T_2^n is a function of X_2^n, (c) follows from $I(T^n;Y_2^n)\le H(T^n)=I(T^n;Y_1^n,X_1^n)=I(T_1^n;Y_1^n|X_1^n)$, using the assumption that $H(T^n|Y_1^n,X_1^n)=0$ and T_1^n is independent of X_1^n. Finally, (d) uses that Y_{1i} is determined by (X_{1i},T_i) and the memorylessness of the channel $p_{Y_2|X_1,X_2}$. □

3.2. **Another class of interference channels.**

Theorem 7. *Consider an interference channel of the form* $p(y_1,y_2|x_1,x_2)=p(y_2|x_2)p(y_1|x_1,y_2)$. *Set*

$$\begin{aligned} C_1 &= \max_{p(x_1),x_2} I(X_1;Y_1|X_2=x_2), \\ C_2 &= \max_{p(x_2),x_1} I(X_2;Y_2|X_1=x_1). \end{aligned}$$

Assume that

$$C_1=\max_{p(x_1),p(x_2)} I(X_1,X_2;Y_1).$$

Further, assume that for any $p(x_1)p(x_2)$ *we have*

$$\frac{I(X_2;Y_2)}{C_2}\le\frac{I(X_2;Y_1|X_1)}{C_1}.$$

Then, the capacity region of this interference channel is the set of non-negative (R_1,R_2) *satisfying*

$$\frac{R_1}{C_1}+\frac{R_2}{C_2}\le 1.$$

In other words, the time-division strategy is optimal.

Proof. Set $\lambda = C_1/C_2$. Observe that

$$\begin{aligned} n(R_1 + \lambda R_2) &\leq I(X_1^n; Y_1^n) + \lambda I(X_2^n; Y_2^n) \\ &= I(X_1^n; Y_1^n) + \lambda \sum_{i=1}^{n} I(X_{2i}; Y_{2i}|Y_2^{i-1}) \\ &\leq \sum_{i=1}^{n} I(X_1^n, Y_1^{i-1}; Y_{1i}) + \lambda I(X_{2i}; Y_{2i}|Y_2^{i-1}) \\ &\leq \sum_{i=1}^{n} I(X_1^n, Y_1^{i-1}, Y_2^{i-1}; Y_{1i}) + \lambda I(X_{2i}; Y_{2i}|Y_2^{i-1}) \\ &= \sum_{i=1}^{n} I(X_{1i}, Y_2^{i-1}; Y_{1i}) + \lambda I(X_{2i}; Y_{2i}|Y_2^{i-1}). \end{aligned}$$

The last equality is due to the fact that

$$(X_1^n, Y_1^{i-1}) \to (X_{1i}, Y_2^{i-1}) \to Y_{1i}$$

forms a Markov chain which follows from the joint decomposition decomposes as $p(x_1^n)p(x_2^n)\prod_i p(y_{2i}|x_{2i})p(y_{1i}|x_{1i}, y_{2i})$.

Let $U_i = Y_2^{i-1}$. Therefore, for some $p(u, x_2)p(x_1)$, we have

$$\begin{aligned} R_1 + \lambda R_2 &\leq I(U, X_1; Y_1) + \frac{C_1}{C_2} I(X_2; Y_2|U) \\ &= I(X_1, X_2; Y_1) - I(X_2; Y_1|X_1, U) + \frac{C_1}{C_2} I(X_2; Y_2|U) \\ &\leq C_1, \end{aligned}$$

where the last inequality follows since $I(X_1, X_2; Y_1) \leq C_1$ and for every u,

$$-I(X_2; Y_1|X_1, U = u) + \frac{C_1}{C_2} I(X_2; Y_2|U = u) \leq 0.$$

□

3.2.1. *An application.* As an application of the above theorem, consider an interference channel with binary inputs X_1 and X_2 taking values in $\{0, 1\}$. We begin with some historical remarks. The class of binary input *deterministic* interference channels was originally studied by Etkin and Ordentlich [9]. They established the capacity region for every such setting, except for the AND-OR setting, i.e. $Y_1 = X_1 \wedge X_2$, and $Y_2 = X_1 \vee X_2$. In particular, they solved the problem when $Y_2 = X_2$ and $Y_1 = X_1 \oplus X_2$ where $\oplus$ denote the sum (modulo 2). In this paper, we consider a non-deterministic variant of this channel by passing the channel outputs through erasure channels. This construction is motivated by a similar construction in [4] in the context of broadcast channels (in [4], outputs of the Blackwell broadcast channel are passed through an erasure channel).

We now give a formal definition. Let $\hat{Y}_1 = X_1 \oplus X_2$ and $\hat{Y}_2 = X_2$. The real receivers Y_1 and Y_2 are obtained by passing $\hat{Y}_1$ and $\hat{Y}_2$ through symmetric BECs, with erasure probabilities ϵ_1 and ϵ_2 respectively. We have $C_1 = 1 - \epsilon_1$ and $C_2 = 1 - \epsilon_2$ and the two corner points of the capacity region are $(R_1, R_2) = (1 - \epsilon_1, 0)$ and $(R_1, R_2) = (0, 1 - \epsilon_2)$.

Theorem 8. *If $\epsilon_1 > \epsilon_2$, the capacity region is the time-division region, i.e., the set of non-negative rate pairs (R_1, R_2) satisfying*

$$R_1 + \frac{1 - \epsilon_1}{1 - \epsilon_2} R_2 \leq 1 - \epsilon_1.$$

If $\epsilon_1 \leq \epsilon_2$, the capacity region is the set of (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq 1 - \epsilon_1 \\ R_2 &\leq 1 - \epsilon_2 \\ R_1 + R_2 &\leq 1 - \epsilon_1. \end{aligned}$$

Proof. We consider two cases:

- The first case is when $\epsilon_1 > \epsilon_2$. We need to show that the capacity region is the time-division region. We apply Theorem 7. First, we couple the outputs of the interference channels such that whenever erasure occurs in the second channel, erasure also occurs in the first channel. This allows us to assume that the joint decomposition decomposes as $p(x_1)p(x_2)p(y_2|x_2)p(y_1|x_1, y_2)$. Note that

$$\max_{p(x_1)p(x_2)} I(X_1, X_2; Y_1) = (1 - \epsilon_1) \max_{p(x_1)p(x_2)} I(X_1, X_2; \hat{Y}_1) = 1 - \epsilon_1.$$

Moreover, for any $p(x_1)p(x_2)$ we have $I(X_2; Y_2) = (1-\epsilon_2)H(X_2)$ and $I(X_2; Y_1|X_1) = (1-\epsilon_1)H(X_2|X_1) = (1 - \epsilon_1)H(X_2)$. Thus,

$$\frac{I(X_2; Y_2)}{C_2} \leq \frac{I(X_2; Y_1|X_1)}{C_1}$$

holds with equality.

- Assume that $\epsilon_1 \leq \epsilon_2$. Then observe that $I(X_2; Y_1|X_1) \geq I(X_2; Y_2|X_1)$ for any $p(x_1)p(x_2)$, i.e., strong interference from the second transmitter (see Definition 1). Strong interference from the second transmitter implies that

$$R_1 + R_2 \leq \max_{p(x_1)p(x_2)} I(X_1, X_2; Y_1) = 1 - \epsilon_1.$$

We also have $R_1 \leq C_1 = 1 - \epsilon_1$ and $R_2 \leq C_2 = 1 - \epsilon_2$. This shows that any achievable rate pair must satisfy

$$\begin{aligned} R_1 &\leq 1 - \epsilon_1 \\ R_2 &\leq 1 - \epsilon_2 \\ R_1 + R_2 &\leq 1 - \epsilon_1. \end{aligned}$$

To show that the above region is achievable, take the following auxiliary variables $U_1 = X_1, U_2 = X_2$ in the Han-Kobayashi bound in Theorem 1.

We remark that in this case, the channel described in the statement of the theorem is a Z-interference channel (Y_2 sees no interference), and we have strong interference at receiver Y_1. The capacity region in this case is known as one can effectively provide X_1 for free to receiver Y_2 (thereby potentially increasing the capacity region). Observe that here $I(X_1; Y_2, X_1|X_2) \geq I(X_1; Y_1|X_2)$. Therefore, it becomes a strong interference channel whose capacity region [10] is obtained by employing MAC decoding at Y_1 and Y_2. Note that the constraint on R_2 is $R_2 \leq I(X_2; Y_2, X_1|X_1) = I(X_2; Y_2)$. Therefore, the capacity region in this setting is (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y_1|X_2, Q) \\ R_2 &\leq I(X_2; Y_2) \\ R_1 + R_2 &\leq I(X_1, X_2; Y_1|Q) \end{aligned}$$

for some $p_Q p_{X_1|Q} p_{X_2|Q}$.

□

References

[1] A. El Gamal and Y.-H. Kim, *Network Information Theory.* Cambridge University Press, 2012.

[2] H. Sato, "The capacity of the gaussian interference channel under strong interference (corresp.)," *IEEE Transactions on Information Theory*, vol. 27, no. 6, pp. 786–788, Nov 1981.

[3] M. Costa, "A new entropy power inequality," *IEEE Transactions on Information Theory*, vol. 31, no. 6, pp. 751–760, 1985.

[4] A. Gohari and C. Nair, "Outer bounds for multiuser settings: the auxiliary receiver approach," *IEEE Transactions on Information Theory*, vol. 68, no. 2, pp. 701–736, 2022.

[5] M. H. M. Costa, "On the Gaussian interference channel," *Information Theory, IEEE Transactions on*, vol. 31, no. 5, pp. 607–615, Sep 1985.

[6] Y. Geng and C. Nair, "The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2087–2104, April 2014.

[7] G. Darmois, "Analyse générale des liaisons stochastiques: etude particulière de l'analyse factorielle linéaire," *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, vol. 21, no. 1/2, pp. pp. 2–8, 1953. [Online]. Available: http://www.jstor.org/stable/1401511

[8] V. P. Skitovic, “Linear combinations of independent random variables and the normal distribution law,” *Select Transl Math Stat Probab*, vol. 2, no. 2, pp. 211–228, 1962.

[9] R. H. Etkin and E. Ordentlich, “Analysis of deterministic binary interference channels via a general outer bound,” *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2597–2604, May 2011.

[10] M. H. M. Costa and A. El Gamal, “The capacity region of the discrete memoryless interference channel with strong interference (corresp.),” *Information Theory, IEEE Transactions on*, vol. 33, no. 5, pp. 710–711, 1987.