# Very Weak Interference Channels

Sida Liu, Chandra Nair, Lingxiao Xia
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, NT, Hong Kong S.A.R.
Email:{sida-liu,chandra,lingxiao}@ie.cuhk.edu.hk

***Abstract*—We derive a genie-based outer bound for the sum rate of discrete memoryless interference channels. We define a class of very weak interference channels and study a sub-class called the binary skewed-Z interference channel. We use the genie based outer bound to deduce the sum-capacity in a non-trivial regime of parameters for this sub-class.**

## I. Introduction

The interference channel is a model for communication of two (or more) pairs of transmitters and receivers over a common medium. Each sender wants to send a private message to its intended receiver and one is interested in characterizing the region of rate-pairs that are simultaneously achievable, *i.e.* the capacity region. The characterization of the capacity region is a classical and fundamental open problem in multi-terminal information theory. For some background on this problem and problem definition, please refer to Chapter 6 in [4].

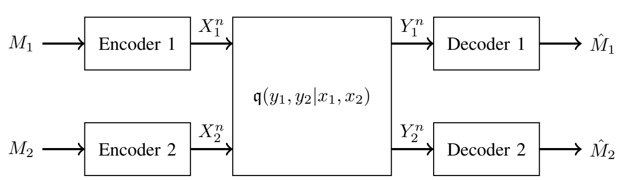



Fig. 1. Discrete memoryless interference channel

A rate pair $(R_1, R_2)$ is said to be achievable if there is a sequence of encoding schemes such that $P_e := Pr\{(M_1, M_2) \neq (\hat{M}_1, \hat{M}_2)\} \to 0$ as $n \to \infty$, when the messages $M_1, M_2$ are distributed uniformly over $[1 : \lfloor 2^{nR_1} \rfloor] \times [1 : \lfloor 2^{nR_2} \rfloor]$. The *capacity region* is the closure of the set of achievable rate pairs $(R_1, R_2)$.

In this paper, we restrict ourselves to maximizing the sum-rate $(R_1 + R_2)$.

## II. Inner and outer bounds for the sum-rate

### A. Inner bound

The sum-capacity of interference channel is not known in general. The best known achievable region is the Han-Kobayashi inner bound [5], [3], which subsumes all other known inner bounds. Performing Fourier-Motzkin elimination on this region will allow us to obtain the corresponding sum-rate inner bound.

**Theorem 1** (Han-Kobayashi sum-rate inner bound)**.** *Any non-negative value $R_1 + R_2$ satisfying the constraints*

$$R_1 + R_2 \leq I(X_1; Y_1|U_2, Q) + I(X_2; Y_2|U_1, Q) \tag{1a}$$
$$R_1 + R_2 \leq I(U_2X_1; Y_1|Q) + I(X_2; Y_2|U_2U_1Q) \tag{1b}$$
$$R_1 + R_2 \leq I(U_1X_2; Y_2|Q) + I(X_1; Y_1|U_2U_1Q) \tag{1c}$$
$$R_1 + R_2 \leq I(U_2X_1; Y_1|U_1Q) + I(U_1X_2; Y_2|U_2Q) \tag{1d}$$

*for some $p(q)p(u_1, x_1|q)p(u_2, x_2|q)$ is achievable.*

There are some outer bounds for the discrete memoryless interference channel

### B. A routine outer bound

**Theorem 2.** *It can be shown*[1] *that any achievable rate pair $(R_1, R_2)$ must satisfy*

$$\begin{aligned} R_1 &\leq \min\{I(U_2X_1; Y_1|Q), I(X_1; Y_1|X_2Q)\} \\ R_2 &\leq \min\{I(U_1X_2; Y_2|Q), I(X_2; Y_2|X_1Q)\} \\ R_1 + R_2 &\leq I(U_2X_1; Y_1|Q) + I(X_2; Y_2|U_2X_1Q) \\ R_1 + R_2 &\leq I(U_1X_2; Y_2|Q) + I(X_1; Y_1|U_1X_2Q), \end{aligned}$$

*for some distribution of the form $p(q, u_1, u_2, x_1, x_2) = p(q)p(u_1, u_2|q)p(x_1|u_1, u_2, q)p(x_2|u_1, u_2, q)$.*

To deduce this outer bound one can use routine arguments along with the following identification for the auxiliaries: $U_{1i} = (X_2^{n\setminus i}, Y_{11}^{i-1}, Y_{2i+1}^n), U_{2i} = (X_1^{n\setminus i}, Y_{11}^{i-1}, Y_{2i+1}^n)$. The first two constraints are straightforward. To obtain the third constraint, observe that (by Fano's inequality)

$$\begin{aligned} &n(R_1 + R_2) - n\epsilon_n \\ \leq& I(X_1^n; Y_1^n) + I(X_2^n; Y_2^n|X_1^n) \\ \overset{(a)}{=}& \sum_{i=1}^n I(X_1^n, Y_{2i+1}^n; Y_{1i}|Y_{11}^{i-1}) + I(X_2^n; Y_{2i}|X_1^n, Y_{2i+1}^n, Y_{11}^{i-1}) \\ \leq& \sum_{i=1}^n I(U_{2i}, X_{1i}; Y_{1i}) + I(X_{2i}; Y_{2i}|U_{2i}, X_{1i}) \end{aligned}$$

where $(a)$ uses Csiszar sum lemma. The fourth constraint follows similarly and the Markov chains are reasonably straightforward to establish using the $d$-separation principle of Bayesian networks.

[1]This outer bound is due to Nair. It is an unpublished result first used as a mid-term question in Network Information Theory class: a version of it was used in Fall '09 and a complete version in Fall '11. In 2012, a similar version of this was independently discovered by other researchers as well.

Applying Fourier-Motzkin elimination on the aforementioned outer bound and eliminating redundant inequalities, we obtain the following sum-rate outer bound for an interference channel.

**Corollary 1** (Sum-rate outer bound)**.** *Any sum-rate $R_1 + R_2$ that is achievable must satisfy the constraints*

$$R_1 + R_2 \leq I(X_1; Y_1|X_2Q) + I(X_2; Y_2|X_1Q) \quad \text{(2a)}$$
$$R_1 + R_2 \leq I(U_2X_1; Y_1|Q) + I(U_1X_2; Y_2|Q) \quad \text{(2b)}$$
$$R_1 + R_2 \leq I(U_2X_1; Y_1|Q) + I(X_2; Y_2|U_2X_1Q) \quad \text{(2c)}$$
$$R_1 + R_2 \leq I(U_1X_2; Y_2|Q) + I(X_1; Y_1|U_1X_2Q) \quad \text{(2d)}$$

*for some distribution of the form $p(q, u_1, u_2, x_1, x_2) = p(q)p(u_1, u_2|q)p(x_1|u_1, u_2, q)p(x_2|u_1, u_2, q)$.*

### *C. Genie based outer bound*

In the scalar Gaussian interference channel it was shown that treating interference as noise is optimal, for sum-capacity, under a certain weak interference condition (see Chapter 6 in [4]). The optimality (or converse) was shown using "genie-aided" receivers. Inspired by this technique, we develop the following outer bound for a general discrete memoryless interference channel. We will then show that this new outer bound helps us determine the sum-capacity for certain new classes of discrete memoryless interference channels, in addition to recovering the Gaussian result.

**Theorem 3.** *Let $T_1, T_2$ be any pair of random variables such that: $p(y_1, t_1|x_1, x_2) = p(t_1|x_1)p(y_1|t_1, x_1, x_2)$, $p(y_2, t_2|x_1, x_2) = p(t_2|x_2)p(y_2|t_2, x_1, x_2)$, and the marginals are consistent with the given channel transition probabilities, i.e. $p(y_1|x_1, x_2) = \mathfrak{q}(y_1|x_1, x_2)$ and $p(y_2|x_1, x_2) = \mathfrak{q}(y_2|x_1, x_2)$. The achievable sum-rate of the discrete memoryless interference channel characterized by $\mathfrak{q}(y_1, y_2|x_1, x_2)$ can be upper bounded as follows:*

$$\begin{aligned} R_1 + R_2 \leq \max_{p_1(x_1)p_2(x_2)} & I(X_1; T_1Y_1) + I(X_2; T_2Y_2) \\ & + \mathfrak{C}[I(X_2; T_2|X_1T_1) - I(X_2; Y_1|T_1X_1)] \\ & - I(X_2; T_2|X_1T_1) + I(X_2; Y_1|T_1X_1) \\ & + \mathfrak{C}[I(X_1; T_1|X_2T_2) - I(X_1; Y_2|T_2X_2)] \\ & - I(X_1; T_1|X_2T_2) + I(X_1; Y_2|T_2X_2), \end{aligned} \quad (3)$$

*where $\mathfrak{C}[I(X_2; T_2|X_1T_1) - I(X_2; Y_1|T_1X_1)]$ denotes the upper concave envelope of the function $I(X_2; T_2|X_1T_1) - I(X_2; Y_1|T_1X_1)$ with respect to product distributions $p_a(x_1)p_b(x_2)$ evaluated at $p_1(x_1)p_2(x_2)$. Similarly the term $\mathfrak{C}[I(X_1; T_1|X_2T_2) - I(X_1; Y_2|T_2X_2)]$ denotes the upper concave envelope of the function $I(X_1; T_1|X_2T_2) - I(X_1; Y_2|T_2X_2)$ with respect to to product distributions $p_a(x_1)p_b(x_2)$ evaluated at $p_1(x_1)p_2(x_2)$.*

*Proof:* Consider a sequence of codebooks such that their decoding error probabilities tend to zero as the block length $n$ tends to infinity. A distribution on the $n$-tuples given by

$$\begin{aligned} & p(m_1, m_2, x_1^n, x_2^n, y_1^n, t_1^n, y_2^n, t_2^n) \\ =& p(m_1, x_1^n)p(m_2, x_2^n) \\ & \prod_{i=1}^n p(t_{1i}|x_{1i})p(y_{1i}|x_{1i}, x_{2i}, t_{1i})p(t_{2i}|x_{2i})p(y_{2i}|x_{1i}, x_{2i}, t_{2i}). \end{aligned}$$

Keep in mind that the capacity only depends on the marginals $\mathfrak{q}(y_1|x_1, x_2)$ and $\mathfrak{q}(y_2|x_1, x_2)$ and the above distribution is consistent with the marginal distributions by the assumptions on $(T_1, T_2)$. One can get an upper bound on sum-rate by following manipulations. The initial part mimics the manipulations in the Gaussian argument as presented in the Appendix of Chapter 6 in [4].

$$\begin{aligned} & n(R_1 + R_2) \\ =& H(M_1) + H(M_2) \\ \leq& I(M_1; Y_1^n) + I(M_2; Y_2^n) + n\epsilon \qquad \text{(by Fano's inequality)} \\ \leq& I(X_1^n; Y_1^n) + I(X_2^n; Y_2^n) + n\epsilon \\ \leq& I(X_1^n; Y_1^nT_1^n) + I(X_2^n; Y_2^nT_2^n) + n\epsilon \\ =& I(X_1^n; T_1^n) + I(X_1^n; Y_1^n|T_1^n) \\ & + I(X_2^n; T_2^n) + I(X_2^n; Y_2^n|T_2^n) + n\epsilon \\ =& H(T_1^n) - H(T_1^n|X_1^n) + H(Y_1^n|T_1^n) - H(Y_1^n|T_1^nX_1^n) \\ +& H(T_2^n) - H(T_2^n|X_2^n) + H(Y_2^n|T_2^n) - H(Y_2^n|T_2^nX_2^n) + n\epsilon. \end{aligned}$$

Consider the term $H(T_1^n) - H(Y_2^n|X_2^nT_2^n)$. Note that

$$\begin{aligned} & H(T_1^n) - H(Y_2^n|X_2^nT_2^n) \\ =& H(T_1^n|T_2^nX_2^n) - H(Y_2^n|X_2^nT_2^n) \\ & \qquad \text{(since } T_1^n \text{ is independent of } (T_2^n, X_2^n)) \\ =& \sum_i H(T_{1i}|T_1^{i-1}T_2^nX_2^n) - H(Y_{2i}|Y_{2,i+1}^nX_2^nT_2^n) \\ =& \sum_i H(T_{1i}|Y_{2,i+1}^nT_1^{i-1}T_2^nX_2^n) - H(Y_{2i}|T_1^{i-1}Y_{2,i+1}^nX_2^nT_2^n) \\ & \qquad \text{(Csiszar-sum lemma)} \\ =& \sum_i H(T_{1i}|U_iX_{2i}T_{2i}) - H(Y_{2i}|U_iX_{2i}T_{2i}). \\ & \qquad ( U_i := (Y_{2,i+1}^n, T_1^{i-1}, T_2^{n\setminus i}, X_2^{n\setminus i})) \end{aligned}$$

One can verify that $X_{1i} \to U_i \to X_{2i}$ in the Appendix.

Similarly

$$\begin{aligned} & H(T_2^n) - H(Y_1^n|X_1^nT_1^n) \\ =& \sum_i H(T_{2i}|V_iX_{1i}T_{1i}) - H(Y_{1i}|V_iX_{1i}T_{1i}) \end{aligned}$$

where $V_i = (Y_{1,i+1}^n, T_2^{i-1}, T_1^{n\setminus i}, X_1^{n\setminus i})$ and $X_{1i} \to V_i \to X_{2i}$.

By the $n$-tuple distribution,

$$H(T_1^n|X_1^n) = \sum_{i=1}^n H(T_{1i}|X_{1i}X_1^{n\setminus i}T_1^{i-1}) = \sum_{i=1}^n H(T_{1i}|X_{1i}),$$
$$H(T_2^n|X_2^n) = \sum_{i=1}^n H(T_{2i}|X_{2i}X_2^{n\setminus i}T_2^{i-1}) = \sum_{i=1}^n H(T_{2i}|X_{2i}).$$

From the chain rule and that conditioning reduces entropy we also obtain

$$H(Y_1^n|T_1^n) \leq \sum_{i=1}^{n} H(Y_{1i}|T_{1i}),$$
$$H(Y_2^n|T_2^n) \leq \sum_{i=1}^{n} H(Y_{2i}|T_{2i}).$$

Combining the above arguments, we obtain using routine manipulations that

$$\begin{aligned}
&n(R_1+R_2)\\
\leq& H(T_1^n) - H(T_1^n|X_1^n) + H(Y_1^n|T_1^n) - H(Y_1^n|T_1^n X_1^n)\\
&+ H(T_2^n) - H(T_2^n|X_2^n) + H(Y_2^n|T_2^n) - H(Y_2^n|T_2^n X_2^n) + n\epsilon\\
\leq& \sum_i H(T_{2i}|V_i X_{1i} T_{1i}) - H(Y_{1i}|V_i X_{1i} T_{1i})\\
&\quad - H(T_{1i}|X_{1i}) + H(Y_{1i}|T_{1i})\\
&\quad + H(T_{1i}|U_i X_{2i} T_{2i}) - H(Y_{2i}|U_i X_{2i} T_{2i})\\
&\quad - H(T_{2i}|X_{2i}) + H(Y_{2i}|T_{2i}) + n\epsilon\\
=& \sum_i I(X_{2i};T_{2i}|V_i X_{1i} T_{1i}) + I(V_i X_{1i}; Y_{1i}|T_{1i})\\
&\quad + I(X_{1i};T_{1i}|U_i X_{2i} T_{2i}) + I(U_i X_{2i}; Y_{2i}|T_{2i}) + n\epsilon\\
=& \sum_i I(X_{2i};T_{2i}|X_{1i} T_{1i}) - I(V_i;T_{2i}|X_{1i}T_{1i})\\
&\qquad \text{(since } I(V_i X_{2i};T_{2i}|X_{1i}T_{1i}) = I(X_{2i};T_{2i}|X_{1i}T_{1i})\text{ )}\\
&\quad + I(X_{1i};Y_{1i}|T_{1i}) + I(V_i;Y_{1i}|T_{1i}X_{1i})\\
&\quad + I(X_{1i};T_{1i}|X_{2i}T_{2i}) - I(U_i;T_{1i}|X_{2i}T_{2i})\\
&\qquad \text{(since } I(U_i X_{1i};T_{1i}|X_{2i}T_{2i}) = I(X_{1i};T_{1i}|X_{2i}T_{2i})\text{ )}\\
&\quad + I(X_{2i};Y_{2i}|T_{2i}) + I(U_i;Y_{2i}|T_{2i}X_{2i}) + n\epsilon\\
=& \sum_i I(X_{2i};T_{2i}) - I(V_i;T_{2i}|X_{1i}T_{1i})\\
&\quad + I(X_{1i};Y_{1i}|T_{1i}) + I(V_i;Y_{1i}|T_{1i}X_{1i})\\
&\quad + I(X_{1i};T_{1i}) - I(U_i;T_{1i}|X_{2i}T_{2i})\\
&\quad + I(X_{2i};Y_{2i}|T_{2i}) + I(U_i;Y_{2i}|T_{2i}X_{2i}) + n\epsilon\\
&\qquad \text{(since } (X_1,T_1) \text{ and } (X_2,T_2) \text{ are independent.)}\\
=& \sum_i I(X_{1i};T_{1i}Y_{1i}) + I(X_{2i};T_{2i}Y_{2i})\\
&\quad - I(V_i;T_{2i}|X_{1i}T_{1i}) + I(V_i;Y_{1i}|T_{1i}X_{1i})\\
&\quad - I(U_i;T_{1i}|X_{2i}T_{2i}) + I(U_i;Y_{2i}|T_{2i}X_{2i}) + n\epsilon
\end{aligned}$$

Now since $V_i \to (X_{1i}, T_{1i}, X_{2i}) \to (Y_{1i}, T_{2i})$ and $U_i \to (X_{1i}, X_{2i}, T_{2i}) \to (Y_{2i}, T_{1i})$, one can rewrite the above as

$$\begin{aligned}
&n(R_1+R_2)\\
\leq& \sum_i I(X_{1i};T_{1i}Y_{1i}) + I(X_{2i};T_{2i}Y_{2i})\\
&\quad - I(X_{2i};T_{2i}|X_{1i}T_{1i}) + I(X_{2i};Y_{1i}|T_{1i}X_{1i})\\
&\quad + I(X_{2i};T_{2i}|V_i, X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|V_i, T_{1i}X_{1i})\\
&\quad - I(X_{1i};T_{1i}|X_{2i}T_{2i}) + I(X_{1i};Y_{2i}|T_{2i}X_{2i})\\
&\quad + I(X_{1i};T_{1i}|U_i, X_{2i}T_{2i}) - I(X_{1i};Y_{2i}|U_i, T_{2i}X_{2i}) + n\epsilon\\
\leq& \sum_i I(X_{1i};T_{1i}Y_{1i}) + I(X_{2i};T_{2i}Y_{2i}) + n\epsilon\\
&\quad - I(X_{2i};T_{2i}|X_{1i}T_{1i}) + I(X_{2i};Y_{1i}|T_{1i}X_{1i})\\
&\quad + \mathfrak{C}[I(X_{2i};T_{2i}|X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|T_{1i}X_{1i})]\\
&\quad - I(X_{1i};T_{1i}|X_{2i}T_{2i}) + I(X_{1i};Y_{2i}|T_{2i}X_{2i})\\
&\quad + \mathfrak{C}[I(X_{1i};T_{1i}|X_{2i}T_{2i}) - I(X_{1i};Y_{2i}|T_{2i}X_{2i})].
\end{aligned}$$

Here $\mathfrak{C}[I(X_{2i};T_{2i}|X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|T_{1i}X_{1i})]$ represents an upper concave envelope of the function $I(X_{2i};T_{2i}|X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|T_{1i}X_{1i})$ defined on the space of distributions $p_1(x_1)p_2(x_2)$. It is easy to see from the definition of the upper concave envelope that

$$\begin{aligned}
&\mathfrak{C}[I(X_{2i};T_{2i}|X_{1i}T_{1i}) - I(X_{2i};Y_{1i}|T_{1i}X_{1i})]\\
=& \sup_{\substack{U: X_{1i} \to U \to X_{2i} \\ U \to (X_{1i}, X_{2i}) \to (Y_{1i}, T_{2i}, T_{1i})}} I(X_{1i};T_{1i}|U, X_{2i}T_{2i})\\
&\qquad - I(X_{1i};Y_{2i}|U, T_{2i}X_{2i}).
\end{aligned}$$

By Bunt's extension of Caratheodory's theorem indeed it suffices to consider $U$ such that $|\mathcal{U}| \leq |\mathcal{X}_1||\mathcal{X}_2|$ to compute the upper concave envelope.

Thus for any valid choice of genies $T_1, T_2$, we obtain an outer bound to the sum-rate given by

$$\begin{aligned}
&R_1 + R_2\\
\leq& \max_{p_1(x_1)p_2(x_2)} I(X_1;T_1Y_1) + I(X_2;T_2Y_2)\\
&\quad + \mathfrak{C}[I(X_2;T_2|X_1T_1) - I(X_2;Y_1|T_1X_1)]\\
&\quad - I(X_2;T_2|X_1T_1) + I(X_2;Y_1|T_1X_1)\\
&\quad + \mathfrak{C}[I(X_1;T_1|X_2T_2) - I(X_1;Y_2|T_2X_2)]\\
&\quad - I(X_1;T_1|X_2T_2) + I(X_1;Y_2|T_2X_2)
\end{aligned} \tag{4}$$

■

**Remark 1.** *The following observations are worth noting.*

*(a) Since every valid pair $T_1, T_2$ (or genies) yields an outer bound, one minimizes the above expression over the choice of valid genies to obtain the best genie based outer bound for the sum-rate. However, since every pair of valid genies yields an outer bound, it is not necessary to provide a cardinality bound on the size of the genie that one needs to consider to make the above region computable.*

*(b) The above genie based outer bound recovers the known result in the scalar Gaussian weak interference regime. Useful genies [1], [8], [7] turn out to be choices of $T_1, T_2$ so that the functions $I(X_2;T_2|X_1T_1) - I(X_2;Y_1|T_1X_1)$ and $I(X_1;T_1|X_2T_2) - I(X_1;Y_2|T_2X_2)$ become concave in $p_2(x_2)$ and $p_1(x_1)$ respectively. For such genies observe that the outer bound reduces to*

$$R_1 + R_2 \leq \max_{p_1(x_1)p_2(x_2)} I(X_1;T_1Y_1) + I(X_2;T_2Y_2),$$

*since the concave envelope of a concave function is itself. The maximizing distributions $(X_{1*}, X_{2*})$ can be shown to be Gaussian by an application of EPI.*

*Within this class of genies where Gaussian signaling is optimal, smart genies [1] ensure that $X_{i*} \to Y_i \to T_i, i = 1, 2$ becomes Markov. Therefore the presence of useful and smart genies reduces the upper bound to*

$$R_1 + R_2 \leq I(X_{1*}; Y_{1*}) + I(X_{2*}; Y_{2*}),$$

*which is achievable by treating interference as noise.*

(c) *Concave envelopes are just a compact way of representing maximizations over auxiliary random variables.*

## III. Very weak interference class of interference channels

In this section we define the very weak interference class, bearing in mind that our interest is in computing the sum-capacity. Our definition (and nomenclature) is motivated in part by the definition of very strong interference channel [2] presented below.

**Definition 1.** *A DM-IC is said to have very strong interference if*

$$I(X_1; Y_1|X_2) \leq I(X_1; Y_2),$$
$$I(X_2; Y_2|X_1) \leq I(X_2; Y_1)$$

*for all $p_1(x_1)p_2(x_2)$.*

**Remark 2.** *In layman's terms a phrasing of the definition is the following: If the interference at the unintended receiver is so strong that one can decode the interfering signal treating ones own signal as noise at a higher rate than the rate at which the true receiver can decode its intended signal even if some genie provides the interfering signal, then the interference is said to be very strong. The optimal strategy indeed turns out to be to decode the interfering signal first and then decodes ones intended signal.*

In a very weak interference setting one expects the intended receiver to treat the interference signal as noise. Additionally, the true receiver should not even try to decode any part of the interfering signal. Motivated by this intuition, we make the following definition.

**Definition 2.** *A discrete memoryless interference channel characterized by the transition matrix $\mathfrak{q}(y_1, y_2|x_1, x_2)$ is called a very weak interference channel if for every pair of auxiliaries $(U_1, U_2)$ such that the joint probability distribution takes the form $p_1(u_1, x_1)p_2(u_2, x_2)\mathfrak{q}(y_1, y_2|x_1, x_2)$ the following inequalities hold:*

$$I(U_1; Y_1) \geq I(U_1; Y_2|X_2)$$
$$I(U_2; Y_2) \geq I(U_2; Y_1|X_1). \tag{5}$$

**Remark 3.** *The following remarks capture some of the intuition as well as limitations of the above definition of very weak interference channels. It would be nice to formally prove this in the sense of [6]. Such a formal proof seems currently out of reach.*

1) *The term $I(U_1; Y_1)$ captures the rate of information from $U_1$ (a part of $X_1$ or a cloud centre among $X_1^n$ sequences) to $Y_1$ when $Y_1$ tries to decode $U_1$ while treating the rest as noise. However, the receiver $Y_1$ could do some interference cancellation of part of $X_2$ before decoding $U_1$; hence this is an underestimate of the information rate from $U_1$ to $Y_1$.*
*The term $I(U_1; Y_2|X_2)$ captures the rate of information from $U_1$ to receiver $Y_2$, after $Y_2$ has (magically) cleaned any effect from $X_2$. This is the maximum rate from $U_1$ that receiver $Y_2$ can hope to decode.*
*Thus the direction of the inequality states that if $U_1$ (part of $X_1$) is to be decoded at $Y_2$ then this imposes a penalty on the rate from $U_1$ to $Y_1$ even under the most favorable (unfavorable) decoding scenario at $Y_2$ ($Y_1$). Thus if one is interested in maximizing $R_1 + R_2$ then one would expect that $Y_2$ should not attempt to decode any part of $X_1$.*
2) *Note that if one is interested in optimizing $\lambda R_1 + R_2$, $\lambda \neq 1$, then one must use a different criterion than the one given above to expect treating interference as noise to remain optimal. In particular if one were to maximize $R_2$, then as long as channels do not have deterministic components, one expects that to obtain $R_2 = I(X_2; Y_2|X_1)$ the receiver $Y_2$ must end up decoding $X_1$, so treating interference as noise may never be optimal.*

**Proposition 1.** *The conditions given in* (5) *are equivalent to the following conditions: for a fixed $p_2(x_2)$ the function $I(X_1; Y_1) - I(X_1; Y_2|X_2)$ is concave in $p_1(x_1)$ and for a fixed $p_1(x_1)$ the function $I(X_2; Y_2) - I(X_2; Y_1|X_1)$ is concave in $p_2(x_2)$.*

*Proof:* Since $U_1 \to X_1 \to (X_2, Y_1, Y_2)$ is Markov observe that

$$I(U_1; Y_1) \geq I(U_1; Y_2|X_2) \iff$$
$$I(X_1; Y_1) - I(X_1; Y_2|X_2) \geq I(X_1; Y_1|U_1) - I(X_1; Y_2|U_1 X_2).$$

The right hand side is clearly equivalent to concavity w.r.t. $p_1(x_1)$. ∎

**Proposition 2.** *Let $SR_{HK}(\mathfrak{q})$ denote the maximum sum-rate achievable using the Han-Kobayashi encoding strategy. Under the very weak interference channel definition in* (5)*, the Han-Kobayashi sum-rate reduces to*

$$SR_{HK}(\mathfrak{q}) = \max_{p_1(x_1)p_2(x_2)} I(X_1; Y_1) + I(X_2; Y_2).$$

*Proof:* Clearly by setting $Q = U_1 = U_2 = 0$ the trivial random variable (i.e. by treating interference as noise) one can indeed achieve the above sum-rate using the Han-Kobayashi scheme.

To observe the reverse direction consider equation (1d) and note the following

$$\begin{aligned}
&I(U_2 X_1; Y_1|U_1 Q) + I(U_1 X_2; Y_2|U_2 Q) \\
&\overset{(a)}{=} I(U_2 X_1; Y_1|Q) - I(U_1; Y_1|Q) \\
&\quad + I(U_1 X_2; Y_2 Q) - I(U_2; Y_2|Q) \\
&= I(X_1; Y_1|Q) + I(U_2; Y_1|X_1 Q) - I(U_2; Y_2|Q) \\
&\quad + I(X_2; Y_2|Q) - I(U_1; Y_1|Q) + I(U_1; Y_2|X_2 Q) \\
&\overset{(b)}{\leq} I(X_1; Y_1|Q) + I(X_2; Y_2|Q).
\end{aligned}$$

Here $(a)$ is a consequence of the Markov chains $U_1 \to X_1 \to (U_2, X_2, Y_1, Y_2)$ and $U_2 \to X_2 \to (U_1, X_1, Y_1, Y_2)$ which hold conditioned on $Q = q$. Inequality $(b)$ is an immediate consequence of our definition of very weak interference channel. Since the average over $Q$ is dominated by the maximum value the lemma is established. ∎

## IV. Binary skewed-Z interference channel

In this section we introduce a class of channels that satisfies the very weak interference condition for a certain set of parameters. We focus on the sum-rate capacity of this class of channels under very weak interference for the rest of the article.

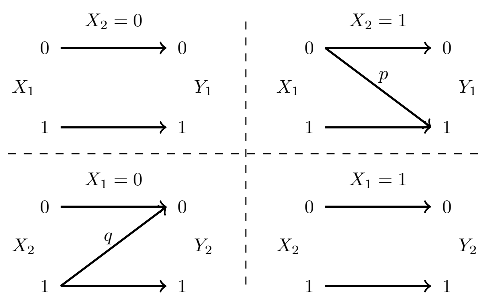



Fig. 2. Binary skewed-Z interference channel (BSZIC)

Figure 2 depicts the transition probabilities of the direct channels for different values of interfering signal. We call such a channel to be *Binary Skewed-Z Interference Channel (BSZIC)*.

**Proposition 3.** *The binary skewed-Z interference channel shown in Figure 2 is a very weak interference channel if and only if* $0 \le p + q \le 1$.

*Proof:* From Proposition 1, it suffices to determine the conditions under which $I(X_1;Y_1) - I(X_1;Y_2|X_2)$ is concave in $p_1(x_1)$ for all fixed $p_2(x_2)$. Let $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ denote the binary entropy function. Let $\mathrm{P}(X_2 = 0) = a$ and $\mathrm{P}(X_1 = 0) = x$. We need to determine the values of $p, q \in [0,1]$ with which $I(X_1;Y_1) - I(X_1;Y_2|X_2)$ is concave in $x$ for all $a \in [0,1]$.

$$\begin{aligned} &I(X_1;Y_1) - I(X_1;Y_2|X_2) \\ =&H(x(1-\bar{a}p)) - xH(1-\bar{a}p) - \bar{a}H(xq) + \bar{a}xH(q), \end{aligned}$$

where $\bar{a} = 1 - a$. Note that the second and the last terms are linear in $x$. After taking second derivative, one could see that the concavity of the above expression w.r.t $x$ is equivalent to showing that

$$\frac{1-\bar{a}p}{1-x(1-\bar{a}p)} \ge \frac{\bar{a}q}{1-xq},$$
$$i.e. \quad (1-\bar{a}p)(1-xq) \ge \bar{a}q(1-x(1-\bar{a}p)).$$

The above condition must hold for every $x \in [0,1]$. Since both sides of the inequality are linear in $x$, it suffices to verify only at $x = 0$ and $x = 1$. Substituting, we obtain the following two conditions, respectively.

$$\begin{cases} 1-\bar{a}p \ge \bar{a}q, \\ (1-\bar{a}p)(1-q) \ge pq\bar{a}^2. \end{cases}$$

Both conditions have to be satisfied at the same time for all $a \in [0,1]$. It is easy to check that this is equivalent to $p+q \le 1$. ∎

**Remark 4.** *We are not able to isolate any non-trivial subset of parameters in the scalar Gaussian interference channel that satisfies the very weak interference condition.*

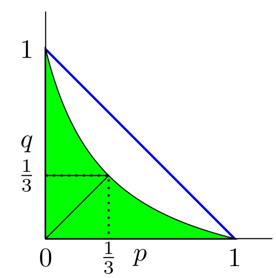



Fig. 3. Regime of parameters where the sum-capacity is established for the Skewed-Z interference channel

**Theorem 4.** *Treating interference as noise is sum-rate optimal for BSZIC with channel parameters* $(p,q)$ *satisfying*

$$\begin{aligned} &0 \le p \le \frac{1}{3}, & & & &0 \le q \le \frac{1}{3}, \\ &p \le q \le \frac{1-p}{1+3p} & &\textit{or} & &q \le p \le \frac{1-q}{1+3q} \end{aligned}$$

*The regime of parameters (as a subset of the weak-interference regime) is shown in Figure 3.*

*Proof:*

In the green region of Figure 3, there is a valid choice of genie $T_1$ , $T_2$ such that

- $X_{i*} \to Y_i \to T_i, i = 1,2$
- The functions $I(X_2;T_2|X_1T_1) - I(X_2;Y_1|T_1X_1)$ and $I(X_1;T_1|X_2T_2) - I(X_1;Y_2|T_2X_2)$ become concave in $p_2(x_2)$ and $p_1(x_1)$ respectively.

This would then imply immediately that the equations (3) reduce to

$$\max_{p_1(x_1)p_2(x_2)} I(x_1;Y_1) + I(X_2;Y_2)$$

which is achievable, hence establishing the sum-capacity.

Let $x = Pr(X_1 = 0)$, $y = Pr(X_2 = 1)$. Consider binary (say $\mathcal{T}_i = \{0,1\}$) genies $T_1$, $T_2$ with the following joint distribution:

| $X_1$ | $X_2$ | $Y_1$ | $T_1$ | Probability |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | $x(1-y)(1-p)a$ |
| 0 | 1 | 0 | 0 | $xy(1-p)a$ |
| 0 | 0 | 0 | 1 | $x(1-y)((1-p)(1-a)+p)$ |
| 1 | 0 | 1 | 1 | $(1-x)(1-y)$ |
| 0 | 1 | 0 | 1 | $xy(1-p)(1-a)$ |
| 0 | 1 | 1 | 1 | $xyp$ |
| 1 | 1 | 1 | 1 | $(1-x)y.$ |

| $X_1$ | $X_2$ | $Y_2$ | $T_2$ | Probability |
|---|---|---|---|---|
| 1 | 1 | 1 | 0 | $(1-x)y(1-q)c$ |
| 0 | 1 | 1 | 0 | $xy(1-q)c$ |
| 1 | 1 | 1 | 1 | $(1-x)y((1-q)(1-c)+q)$ |
| 0 | 1 | 0 | 1 | $xyq$ |
| 1 | 0 | 0 | 1 | $(1-x)(1-y)$ |
| 0 | 1 | 1 | 1 | $xy(1-q)(1-c)$ |
| 0 | 0 | 0 | 1 | $x(1-y).$ |

Parameters $0<a<1$, $0<c<1$ will be determined later. It is easy to check this distribution has the laws $p(y_1|x_1,x_2)$ which are consistent with the channel. So this setting gives a valid genie.

Below we check the Markov chains $X_i \to Y_i \to T_i, i=1,2$

$$\begin{cases} p(T_1=0|X_1=0,Y_1=0)=p(T_1=0|Y_1=0) \\ \qquad =\frac{(1-p)a}{1-y+y(1-p)} \\ p(T_1=0|X_1=0,Y_1=1)=p(T_1=0|X_1=1,Y_1=1) \\ \qquad =p(T_1=0|Y_1=1)=0 \end{cases}$$

$$\begin{cases} p(T_2=0|X_2=1,Y_2=1)=p(T_2=0|Y_2=1) \\ \qquad =\frac{(1-q)c}{1-x+x(1-q)} \\ p(T_2=0|X_2=0,Y_2=0)=p(T_2=0|X_2=1,Y_2=0) \\ \qquad =p(T_2=0|Y_2=0)=0 \end{cases}$$

Second, we show $I(X_2;T_2|X_1T_1)-I(X_2;Y_1|T_1X_1)$ is concave in $p_2(x_2)$ for any distribution of $X_2$:

Define

$$f(x,y):=\big(I(X_2;T_2|X_1T_1)-I(X_2;Y_1|T_1X_1)\big)|_{\mathrm{P}(X_1=0)=x,\mathrm{P}(X_2=1)=y}.$$

For simplicity of notation, for a generic variable $x\in[0,1]$, let $\bar{x}=1-x$ and $L(x)=-x\log_2 x$. Then

$$\begin{aligned} &f(x,y)\\ =&L(y\bar{q}c)-yL(\bar{q}c)+L(\bar{y}+y(\bar{q}\bar{c}+q))-yL(\bar{q}\bar{c}+q)\\ &-(xp+x\bar{p}\bar{a})L\left(\frac{yp}{p+\bar{p}\bar{a}}\right)-x(p+\bar{p}\bar{a})L\left(\frac{\bar{y}p+\bar{p}\bar{a}}{p+\bar{p}\bar{a}}\right)\\ &+xy(p+\bar{p}\bar{a})L\left(\frac{p}{p+\bar{p}\bar{a}}\right)+xy(pb+\bar{p}\bar{a})L\left(\frac{\bar{p}\bar{a}}{p+\bar{p}\bar{a}}\right). \end{aligned}$$

It remains to show $\frac{\partial^2 f}{\partial y^2}\le 0$.

$$\begin{aligned} &\frac{\partial^2 f}{\partial y^2}\\ =&-\frac{\bar{q}c}{y}-\frac{\bar{q}^2c^2}{1-y\bar{q}c}+\frac{xp}{y}+\frac{xp^2}{\bar{y}p+\bar{p}\bar{a}}\\ \le&-\frac{\bar{q}c}{y}-\frac{\bar{q}^2c^2}{1-y\bar{q}c}+\frac{p}{y}+\frac{p^2}{\bar{y}p+\bar{p}\bar{a}}\\ =&-\frac{\bar{q}c}{y(1-y\bar{q}c)}+\frac{p^2+p\bar{p}\bar{a}}{y(\bar{y}p+\bar{p}\bar{a})}\\ =&-\frac{1}{y}\left(\frac{\bar{q}c}{1-y\bar{q}c}-\frac{p^2+p\bar{p}\bar{a}}{\bar{y}p+\bar{p}\bar{a}}\right)\\ =&-\frac{1}{y}\left(\frac{(\bar{p}\bar{a}+p-1)p\bar{q}cy+\bar{q}cp+\bar{q}\bar{p}c\bar{a}-p^2-p\bar{p}\bar{a}}{(1-y\bar{q}c)(\bar{y}p+\bar{p}\bar{a})}\right)\\ \le&-\frac{1}{y}\left(\frac{(\bar{p}\bar{a}+p-1)p\bar{q}c+\bar{q}cp+\bar{q}\bar{p}c\bar{a}-p^2-p\bar{p}\bar{a}}{(1-y\bar{q}c)(\bar{y}p+\bar{p}\bar{a})}\right)\\ =&-\frac{1}{y}\left(\frac{(\bar{p}\bar{a}p+p^2+\bar{p}\bar{a})\bar{q}c-p^2-p\bar{p}\bar{a}}{(1-y\bar{q}c)(\bar{y}p+\bar{p}\bar{a})}\right) \end{aligned}$$

We choose

$$c=\frac{p^2+p\bar{p}\bar{a}}{\bar{q}(p^2+p\bar{p}\bar{a}+\bar{p}\bar{a})}$$

where $a$ is determined later and the validity of $c$ will be shown after $a$ is determined. Then we have

$$\frac{\partial^2 f}{\partial y^2}\le 0$$

Third, we show that $I(X_1;T_1|X_2T_2)-I(X_1;Y_2|T_2X_2)$ is concave in $p_1(x_1)$. We use similar approach:

Define

$$\begin{aligned} &\tilde{f}(x,y)\\ :=&\big(I(X_1;T_1|X_2T_2)-I(X_1;Y_2|T_2X_2)\big)|_{\mathrm{P}(X_1=0)=x,\mathrm{P}(X_2=1)=y}. \end{aligned}$$

and compute its second derivative

$$\frac{\partial^2\tilde{f}}{\partial x^2}\le-\frac{1}{x}\left(\frac{(\bar{q}\bar{c}q+q^2+\bar{q}\bar{c})\bar{p}a-q^2-q\bar{q}\bar{c}}{(1-x\bar{p}a)(\bar{x}q+\bar{q}\bar{c})}\right)$$

To make the second derivative less than 0, it suffices to show for some $a \in [0,1]$,

$$
\begin{aligned}
\bar{p}a \geq & \frac{q^2 + q\bar{q}\bar{c}}{q^2 + q\bar{q}\bar{c} + \bar{q}\bar{c}} \\
= & 1 - \frac{\bar{q}\bar{c}}{q^2 + q\bar{q}\bar{c} + \bar{q}\bar{c}} \\
= & 1 - \frac{1}{\frac{q^2}{\bar{q}\bar{c}} + q + 1} \\
= & 1 - \frac{1}{\frac{q^2}{\bar{q} - \frac{p^2 + p\bar{p}\bar{a}}{p^2 + p\bar{p}\bar{a} + \bar{p}\bar{a}}} + q + 1} \\
= & 1 - \frac{1}{\frac{q^2}{\frac{\bar{q}p^2 + \bar{q}p\bar{p}\bar{a} + \bar{q}\bar{p}\bar{a} - p^2 - p\bar{p}\bar{a}}{p^2 + p\bar{p}\bar{a} + \bar{p}\bar{a}}} + q + 1} \\
= & 1 - \frac{1}{\frac{q^2p^2 + q^2p\bar{p}\bar{a} + q^2\bar{p}\bar{a}}{\bar{q}\bar{p}\bar{a} - p^2q - p\bar{p}\bar{a}q} + q + 1} \\
= & 1 - \frac{\bar{q}\bar{p}\bar{a} - p^2q - p\bar{p}\bar{a}q}{q^2\bar{p}\bar{a} + q\bar{q}\bar{p}\bar{a} + \bar{q}\bar{p}\bar{a} - p^2q - p\bar{p}\bar{a}q} \\
= & \frac{q\bar{p}\bar{a}}{q\bar{p}\bar{a} + \bar{q}\bar{p}\bar{a} - p^2q - p\bar{p}\bar{a}q} \\
= & \frac{q\bar{p}\bar{a}}{\bar{p}\bar{a} - p^2q - p\bar{p}\bar{a}q} \\
= & \frac{q\bar{p}\bar{a}}{(1 - pq)\bar{p}\bar{a} - p^2q}
\end{aligned}
$$

or equivalently,

$$
\begin{aligned}
(1-pq)(1-p)a - (1-pq)(1-p)a^2 - p^2qa \geq & q - qa \\
(1-pq)(1-p)a^2 - (1-p)(1+q)a + q \leq & 0.
\end{aligned}
$$

Now let's take $a = \frac{1+q}{2(1-pq)}$. then

$$
\begin{aligned}
a = & \frac{1+q}{2(1-pq)} \\
\stackrel{(*)}{\leq} & \frac{1+q}{2(1 - q\frac{1-q}{1+3q})} \\
= & \frac{(1+q)(1+3q)}{2(1+q)^2} \\
= & \frac{1+3q}{2+2q} \leq 1
\end{aligned}
$$

where (*) comes from the constraints on $p, q$. Then

$$
\begin{aligned}
& (1-pq)(1-p)a^2 - (1-p)(1+q)a + q \\
= & \frac{(1-q)(1+3q)\left(p - \frac{1-q}{1+3q}\right)}{4(1-pq)} \leq 0
\end{aligned}
$$

Hence $\tilde{f}(x)$ is concave.

It remains to check $c \leq 1$.

$$
\begin{aligned}
c = & \frac{p^2 + p\bar{p}\bar{a}}{\bar{q}(p^2 + p\bar{p}\bar{a} + \bar{p}\bar{a})} \\
= & \frac{p^2 + p\bar{p}(1 - \frac{1+q}{2-2pq})}{\bar{q}(p^2 + p\bar{p}\bar{a} + \bar{p}(1 - \frac{1+q}{2-2pq}))} \\
= & \frac{p^2 + p\bar{p}(\frac{1-2pq-q}{2-2pq})}{\bar{q}(p^2 + p\bar{p}\bar{a} + \bar{p}(1 - \frac{1+q}{2-2pq}))} \\
= & \frac{p^2 + \frac{p\bar{p} - 2p^2\bar{p}q - p\bar{p}q}{2-2pq}}{\bar{q}(p^2 + p\bar{p}\bar{a} + \frac{\bar{p} - 2p\bar{p}q - \bar{p}q}{2-2pq})} \\
= & \frac{p^2 + \frac{p\bar{p}\bar{q} - 2p^2\bar{p}q}{2-2pq}}{\bar{q}(p^2 + \frac{p\bar{p}\bar{q} - 2p^2\bar{p}q}{2-2pq} + \frac{\bar{p} - 2p\bar{p}q - \bar{p}q}{2-2pq})} \\
= & \frac{p^2(2-2pq) + p\bar{p}\bar{q} - 2p^2\bar{p}q}{\bar{q}(p^2(2-2pq) + p\bar{p}\bar{q} - 2p^2\bar{p}q + \bar{p} - 2p\bar{p}q - \bar{p}q)} \\
= & \frac{2p^2 - 2p^2q + p\bar{p}\bar{q}}{\bar{q}(p^2(2-2pq) + p\bar{p}\bar{q} - 2p^2\bar{p}q + \bar{p} - 2p\bar{p}q - \bar{p}q)} \\
= & \frac{p\bar{q}(1+p)}{\bar{q}(p\bar{q}(1+p) + \bar{p} - 2p\bar{p}q - \bar{p}q)} \\
= & \frac{p(1+p)}{p\bar{q} + p^2\bar{q} + \bar{p} - 2p\bar{p}q - \bar{p}q} \\
= & \frac{p(1+p)}{\bar{q} + p^2\bar{q} - 2p\bar{p}q} \\
= & \frac{p(1+p)}{1 + p^2 - (1+q^2)q - (2p - 2p^2)q} \\
= & \frac{p(1+p)}{1 + p^2 - (1 + 2p - p^2)q}
\end{aligned}
$$

By the constraints on $p, q$,

$$
\begin{aligned}
& \frac{p(1+p)}{1 + p^2 - (1 + 2p - p^2)q} \\
\leq & \frac{p(1+p)}{1 + p^2 - (1 + 2p - p^2)\frac{1-p}{1+3p}} \\
= & \frac{p(1+p)(1+3p)}{(1+p^2)(1+3p) - (1+2p-p^2)(1-p)} \\
= & \frac{p(1+p)(1+3p)}{1 + 3p + p^2 + 3p^3 - (1 + 2p - p^2 - p - 2p^2 + p^3)} \\
= & \frac{p(1+p)(1+3p)}{1 + 3p + p^2 + 3p^3 - (1 + p - 3p^2 + p^3)} \\
= & \frac{p(1+p)(1+3p)}{2p + 4p^2 + 2p^3} \\
= & \frac{1+3p}{2+2p} \\
= & 1 - \frac{1-p}{2+2p} \\
\leq & 1
\end{aligned}
$$

■

**Remark 5.** *In appendix we also show that the above conditions on $(p, q)$ are necessary for the existence of genies such* *that the difference of mutual information terms are concave and the Markov chain holds.*

### A. More on the genie based outer bound

In this section, we analyze the necessary conditions[2] when the genie based outer bound for the skewed-Z interference channel reduces to the sum-rate yielded by treating interference as noise. Since our setting is a discrete setting we are able to perform a much more exhaustive analysis of the bound than that possible in the Gaussian setting.

For a given (valid) pair of genies $(T_1, T_2)$ consider the sum-rate outer bound given by Theorem 3. Further let $p_1^*(x_1)p_2^*(x_2)$ be a maximizing product distribution (i.e. the product distribution that yields the outer bound for this particular choice of genies). For the expression in (3) to reduce to

$$I(X_1; Y_1) + I(X_2; Y_2)$$

at $p_1^*(x_1)p_2^*(x_2)$, it is easy to see that the following equalities must hold:

$$\begin{aligned}
&I(X_1; T_1|Y_1) = 0,\\
&I(X_2; T_2|Y_2) = 0,\\
&\mathfrak{C}[I(X_2; T_2|X_1T_1) - I(X_2; Y_1|T_1X_1)]\\
&\quad = I(X_2; T_2|X_1T_1) + I(X_2; Y_1|T_1X_1),\\
&\mathfrak{C}[I(X_1; T_1|X_2T_2) - I(X_1; Y_2|T_2X_2)]\\
&\quad = I(X_1; T_1|X_2T_2) + I(X_1; Y_2|T_2X_2).
\end{aligned}$$

However these inequalities need to hold only at the maximizing distribution $p_1^*(x_1)p_2^*(x_2)$. Further if such genies exist, by virtue of the fact that the expression $I(X_1; Y_1) + I(X_2; Y_2)$ at $p_1^*(x_1)p_2^*(x_2)$ yields an outer bound to the sum-rate, it must also hold that $p_1^*(x_1)p_2^*(x_2)$ is also a maximizer of the expression $I(X_1; Y_1)+I(X_2; Y_2)$ over all product distributions (since the maximum of $I(X_1; Y_1)+I(X_2; Y_2)$ is an achievable sum-rate).

We first restrict our attention to genies (taking values in some finite alphabet) such that the Markov chains $X_1 \to Y_1 \to T_1$ and $X_2 \to Y_2 \to T_2$ hold at some distribution $\mathrm{P}(X_1 = 0) = x_*$ and $\mathrm{P}(X_2 = 1) = y_*$. One can easily verify that for the Markov chains to hold, the probability distributions must take the form

| $X_1$ | $X_2$ | $Y_1$ | $T_1$ | Probability |
|---|---|---|---|---|
| 0 | 0 | 0 | $i$ | $x_*(1-y_*)((1-p)a_i + pb_i))$ |
| 1 | 0 | 1 | $i$ | $(1-x_*)(1-y)b_i$ |
| 0 | 1 | 0 | $i$ | $x_*y_*(1-p)a_i$ |
| 0 | 1 | 1 | $i$ | $x_*y_*pb_i$ |
| 1 | 1 | 1 | $i$ | $(1-x_*)y_*b_i$ |

for some $0 \le a_, b_i \le 1$. A similar structure also holds for the distribution of $(X_1, X_2, Y_2, T_2)$. An interesting observation is that if the Markov chain holds for some $x_*, y_* > 0$ then the Markov condition continues to hold for any product distribution. This is a chance observation (peculiar to the Binary skewed-Z interference channel) which greatly simplified our analysis.

[2] Note that the previous result only dealt with the sufficient conditions.

Among the class of genies that satisfy the Markov chain, one is further interested in a subclass for which the upper concave envelopes of the differences of mutual information match the function value at $p_1^*(x_1)p_2^*(x_2)$. To this end, define $f(x, y)$ as

$$I(X_2; T_2|X_1T_1) - I(X_2; Y_1|T_1X_1)|_{\mathrm{P}(X_1=0)=x, \mathrm{P}(X_2=1)=y}.$$

Expanding the terms and noting the linearity in $x$ can express $f(x, y) = (1-x)g_0(y) + xg_1(y)$, where $g_0(y) = f(0, y)$ is a concave function of $y$ and $g_1(y) = f(1, y)$ is in general neither convex nor concave in the entire interval $y \in [0, 1]$.

The following proposition aids in our computation of the upper concave envelope of $f(x, y)$.

**Proposition 4.** *Let $\mathfrak{C}[f](x, y)$ denote the upper concave envelope of $f(x, y)$ over the space of product distributions notated by $\mathrm{P}(X_1 = 0) = x, \mathrm{P}(X_2 = 1) = y$. Then*

$$\mathfrak{C}[f](x, y) = (1-x)\mathfrak{C}[g_0](y) + x\mathfrak{C}[g_1(y)],$$

*where $\mathfrak{C}[g_0](y), \mathfrak{C}[g_1](y)$ denotes the upper concave envelope ofs $g_0(y), g_1(y)$ respectively over $y \in [0, 1]$.*

*Proof:* Consider a maximizing convex combination: i.e. a probability vector $\{\alpha_i\}$ and points $(x_i, y_i) \in [0, 1] \times [0, 1]$ such that $\sum_i \alpha_i f(x_i, y_i) = \mathfrak{C}[f](x, y)$. We know that

$$\sum_i \alpha_i x_i y_i = xy, \sum_i \alpha_i x_i = x, \sum_i \alpha_i y_i = y.$$

Obtain a new convex combination as follows: with probability $\alpha_i(1-x_i)$ choose $(0, y_i)$ and with probability $\alpha_i x_i$ choose $(1, y_i)$. Observe that

$$\begin{aligned}
&\sum_i \alpha_i(1-x_i)f(0, y_i) + \alpha_i x_i f(1, y_i)\\
&= \sum_i \alpha_i((1-x_i)f(0, y_i) + x_i f(1, y_i))\\
&= \sum_i \alpha_i f(x_i, y_i) = \mathfrak{C}[f](x, y).
\end{aligned}$$

Since $\sum_i \frac{\alpha_i(1-x_i)}{(1-x)} = 1$ and $\sum_i \frac{\alpha_i(1-x_i)}{(1-x)} y_i = y$ we have $\sum_i \alpha_i(1-x_i)f(0, y_i) \le (1-x)\mathfrak{C}[g_0](y)$. Similarly we have $\alpha_i x_i f(1, y_i) \le x\mathfrak{C}[g_1(y)]$. Thus $\mathfrak{C}[f](x, y) \le (1-x)\mathfrak{C}[g_0](y) + x\mathfrak{C}[g_1(y)]$.

The other direction is immediate as one can always take the convex combination that achieves $\mathfrak{C}[g_0](y)$ and the convex combination that achieves $\mathfrak{C}[g_1](y)$ to obtain a value $(1-x)\mathfrak{C}[g_0](y) + x\mathfrak{C}[g_1(y)]$. ∎

For the binary skewed-$Z$ interference channel, $g_0(y)$ is concave and hence $\mathfrak{C}[g_0](y) = g_0(y)$. We will seek to answer the following question: In the class of genies such that the Markov chain holds, are there genies such that $\mathfrak{C}[g_1(y)] = g_1(y)$ at $y^*$, the maximizing distribution? If the answer is affirmative whenever $p + q \le 1$, then the genie based outer bound will yield the sum-capacity in the entire weak interference regime of parameters. However, we shall see that this is not the case.

*1) Genie approach in an intermediate regime:* We restrict our attention to the symmetric case where $p = q$. When $p = q \le \frac{1}{3}$ we observe that there are genies for which $g_1(y)$ is concave when $y \in [0,1]$.

Now we consider the range $\frac{1}{3} \le p = q \le \frac{1}{2}$. Suppose we restrict ourselves to genies with binary alphabets, then $g_1(y)$ displays an interesting behavior. The function is concave in some interval $[0, \hat{y}]$ and convex in the remainder. Hence the concave envelope of $g_1(y)$ matches the function in the interval $[0, y^\dagger]$ ($y^\dagger \le \hat{y}$) and follows the tangent to the curve $g_1(y)$ (at $y^\dagger$) in the interval $[y^\dagger, 1]$. Here $y^\dagger$ is the unique point in $[0,1]$ such the tangent to the curve $g_1(y)$ at $y^\dagger$ passes through $g_1(1)$ when $y = 1$.

Numerical simulations indicate that there are such genies when $0 \le p = q \le 0.39$. Since we have very explicit expressions, it is not difficult to convert the simulations to a complete argument, but we refrain from doing so because of the following negative result.

**Proposition 5.** *For the binary skewed-Z interference channel when $p = q = \frac{1}{2}$, the genie based outer bound is strictly greater than treating interference as noise inner bound.*

*Proof:*
As before define $f(x,y)$ as

$$I(X_2;T_2|X_1T_1) - I(X_2;Y_1|T_1X_1)|_{\mathrm{P}(X_1=0)=x,\mathrm{P}(X_2=1)=y}.$$

The joint laws are as defined in Table I in the Appendix.

We evaluate $f(x,y)$ as follow. For a generic variable $x \in [0,1]$, let $\bar{x} = 1 - x$ and $L(x) = -x\log_2 x$. Then

$$\begin{aligned} f(x,y) = \sum_i \Big( & L(\bar{y}d_i + y(\bar{q}c_i + qd_i)) - \bar{y}L(d_i) \\ & - yL(\bar{q}c_i + qd_i) - (xpb_i + x\bar{p}a_i)L\left(\frac{ypb_i}{pb_i + \bar{p}a_i}\right) \\ & - x(pb_i + \bar{p}a_i)L\left(\frac{\bar{y}pb_i + \bar{p}a_i}{pb_i + \bar{p}a_i}\right) \\ & + xy(pb_i + \bar{p}a_i)L\left(\frac{pb_i}{pb_i + \bar{p}a_i}\right) \\ & + xy(pb_i + \bar{p}a_i)L\left(\frac{\bar{p}a_i}{pb_i + \bar{p}a_i}\right)\Big). \end{aligned}$$

Split $f(x,y)$ into two functions $g_0(y) = f(0,y)$ and $g_1(y) = f(1,y)$ as in Proposition 4. Then

$$g_0(y) := \sum_i L(\bar{y}d_i + y(\bar{q}c_i + qd_i)) - \bar{y}L(d_i) - yL(\bar{q}c_i + qd_i),$$

$$\begin{aligned} g_1(y) := \sum_i \Big( & L(\bar{y}d_i + y(\bar{q}c_i + qd_i)) - \bar{y}L(d_i) - yL(\bar{q}c_i + qd_i) \\ & - (pb_i + \bar{p}a_i)L\left(\frac{ypb_i}{pb_i + \bar{p}a_i}\right) - (pb_i + \bar{p}a_i)L\left(\frac{\bar{y}pb_i + \bar{p}a_i}{pb_i + \bar{p}a_i}\right) \\ & + y(pb_i + \bar{p}a_i)L\left(\frac{pb_i}{pb_i + \bar{p}a_i}\right) \\ & + y(pb_i + \bar{p}a_i)L\left(\frac{\bar{p}a_i}{pb_i + \bar{p}a_i}\right)\Big). \end{aligned}$$

Setting $p = q = \frac{1}{2}$, compute second derivative of $g_1(y)$

$$\begin{aligned} & \frac{d^2 g_1(y)}{dy^2} \\ =& \sum_i \left(-\frac{(c_i - d_i)^2}{2y(c_i - d_i) + 4d_i} + \frac{b_i}{2y} + \frac{b_i^2}{2\bar{y}b_i + 2a_i}\right) \\ =& -\sum_i \frac{(c_i - d_i)^2}{2y(c_i - d_i) + 4d_i} + \sum_i \frac{b_i}{2y} + \sum_i \frac{yb_i^2}{2y(\bar{y}b_i + a_i)} \\ \ge& -\sum_i \frac{c_i^2 + d_i^2}{2y(c_i - d_i) + 4d_i} + \sum_i \frac{b_i}{2y} + \sum_i \frac{yb_i^2}{2y(\bar{y}b_i + a_i)} \\ =& -\sum_i \frac{c_i^2}{2yc_i - 2yd_i + 4d_i} - \sum_i \frac{d_i^2}{2yc_i - 2yd_i + 4d_i} \\ & + \frac{1}{2y} + \sum_i \frac{yb_i^2}{2y(\bar{y}b_i + a_i)} \\ \ge& -\sum_i \frac{c_i^2}{2yc_i} - \sum_i \frac{d_i^2}{-2yd_i + 4d_i} + \frac{1}{2y} + \sum_i \frac{yb_i^2}{2y(\bar{y}b_i + a_i)} \\ =& -\frac{1}{2y} - \frac{1}{-2y + 4} + \frac{1}{2y} \\ & + \frac{\bar{y} + 1}{2}\left(\sum_i \frac{\bar{y}b_i + a_i}{\bar{y} + 1}\frac{b_i^2}{(\bar{y}b_i + a_i)^2}\right) \\ \overset{(a)}{\ge}& -\frac{1}{-2y + 4} + \frac{\bar{y} + 1}{2}\left(\sum_i \frac{\bar{y}b_i + a_i}{\bar{y} + 1}\frac{b_i}{\bar{y}b_i + a_i}\right)^2 \\ =& -\frac{1}{-2y + 4} + \frac{1}{2(\bar{y} + 1)} \\ =& 0 \end{aligned}$$

where $(a)$ follows since $\mathrm{E}(X^2) \ge \mathrm{E}(X)^2$. Thus $g_1(y)$ is convex in general. The only hope for the outer bound to work would be if $g_1(y)$ was a straight line. So, we next analyze if this is possible,

Note $\frac{d^2 g_1(y)}{dy^2} = 0$ would imply that $c_i d_i = 0$ (for the first inequality to be equality) and $a_i = b_i$ (for the inequality labeled $(a)$ to be an equality).

For the symmetric condition to hold, define $\tilde{f}(x,y)$ as

$$I(X_1;T_1|X_2T_2) - I(X_1;Y_2|T_2X_2)|_{\mathrm{P}(X_1=0)=x,\mathrm{P}(X_2=1)=y}$$

Split $\tilde{f}(x,y)$ in same way as for $f(x,y)$,

$$\tilde{f}(x,y) = (1 - y)\tilde{g}_0(x) + y\tilde{g}_1(x)$$

Computing derivative of $\tilde{g}_1(x)$, we have

$$\frac{d^2\tilde{g}_1(x)}{dx^2} \ge 0$$

with equality holding only *iff* $a_i b_i = 0$ and $c_i = d_i$.

Clearly, both equalities cannot hold at the same time. At least one of $g_1$ and $\tilde{g}_1$ is strictly convex. Therefore, for any $(x,y) \in (0,1)^2$,

$$\begin{aligned} & \mathfrak{C}[f](x,y) + \mathfrak{C}[\tilde{f}](x,y) \\ =& x\mathfrak{C}[g_0](y) + (1 - x)\mathfrak{C}[g_1](y) + y\mathfrak{C}[\tilde{g}_0](x) + (1 - y)\mathfrak{C}[\tilde{g}_1](x) \\ >& xg_0(y) + (1 - x)g_1(y) + y\tilde{g}_0(x) + (1 - y)\tilde{g}_1(x) \\ =& f(x,y) + f_{c_i,d_i,a_i,b_i}(y,x) \end{aligned}$$

■

## Conclusion

We defined the class of very weak interference channels and showed that a subset of parameters of a binary skewed-Z interference channel belongs to this class. We developed a genie based outer bound for the sum-rate of discrete memoryless interference channels. Using this outer bound we showed that treating interference as noise is optimal for a subset of parameters of the binary skewed-Z interference channel in the very weak interference regime. We also showed that the genie based outer bound will not reduce to the sum-rate yielded by interference as noise in the entire very weak interference regime. This work shows that employing genies as a mathematical gadget for proving converses remains largely an unexplored area.


## Acknowledgements

We would like to acknowledge interesting discussions with Yin Zi and Bernd Bandemer.

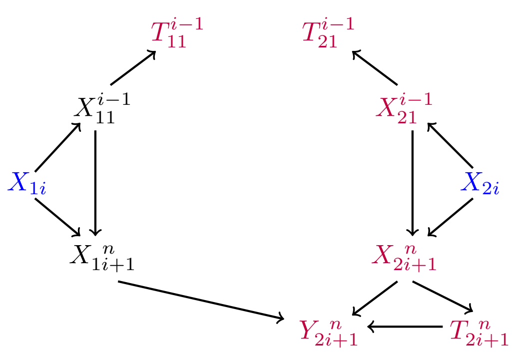

## Appendix

### A. Analysis of dependence in Genie based outer bound

This section aims to verify Markov condition presented in deriving Genie based outer bound

**Lemma 1.** *$X_{1i} \to (Y_{2,i+1}^n, T_1^{i-1}, T_2^{n\setminus i}, X_2^{n\setminus i}) \to X_{2i}$ is Markov.*

*Proof:* Consider a Bayesian network representation of the variables as follows: It is clear that any path from $X_{1i}$ to




$X_{2i}$ is d-separated. Indeed the variable $X_{2i+1}^n$ d-separates the variables into two sets. ■

### B. Necessary condition for Theorem 4

One might doubt if the green region in figure 3 can be improved. As stated in Remark 5, the optimality region must be the region given by Theorem 4, if one insists on genies satisfying the following two conditions.

Markov chain condition:

$X_{i*} \to Y_i \to T_i, i = 1, 2.$

Concavity condition:

The function $I(X_2; T_2|X_1T_1) - I(X_2; Y_1|T_1X_1)$ and $I(X_1; T_1|X_2T_2) - I(X_1; Y_2|T_2X_2)$ become concave in $p_2(x_2)$ and $p_1(x_1)$ respectively.

TABLE I
GENERIC PROBABILITY DISTRIBUTION FOR GENIES THAT SATISFY THE MARKOV CONDITIONS

| $X_1$ | $X_2$ | $Y_1$ | $T_1$ | Probability |
|---|---|---|---|---|
| 0 | 0 | 0 | $i$ | $x(1-y)((1-p)a_i + pb_i))$ |
| 1 | 0 | 1 | $i$ | $(1-x)(1-y)b_i$ |
| 0 | 1 | 0 | $i$ | $xy(1-p)a_i$ |
| 0 | 1 | 1 | $i$ | $xypb_i$ |
| 1 | 1 | 1 | $i$ | $(1-x)yb_i.$ |

| $X_1$ | $X_2$ | $Y_2$ | $T_2$ | Probability |
|---|---|---|---|---|
| 1 | 1 | 1 | $i$ | $(1-x)y((1-q)c_i + qd_i))$ |
| 0 | 1 | 0 | $i$ | $xyqd_i$ |
| 1 | 0 | 0 | $i$ | $(1-x)(1-y)d_i$ |
| 0 | 1 | 1 | $i$ | $xy(1-q)c_i$ |
| 0 | 0 | 0 | $i$ | $x(1-y)d_i.$ |

The catch here is to make genie outer bound optimal, the above two condition need to hold at distribution $p_1^*(x_1)p_2^*(x_2)$, rather than any product distribution. One should be aware that the necessary condition is NOT genie outer bound optimal condition.

Let us restrict our class of genie satisfying Markov condition: $X_{i*} \to Y_i \to T_i, i = 1, 2$. Given that valid genies also satisfy $T_2 \to X_2 \to X_1 \to T_1$, by algebraic manipulations one can verify that the only admissible distributions $p_1(x_1, x_2, y_1, t_1)$ and $p_2(x_1, x_2, y_2, t_2)$ that satisfy the above Markov conditions must be of the form given in Table I.

Here $\{a_i\}, \{b_i\}$ are two generic probability vectors of size $|T_1|$ and $\{c_i\}, \{d_i\}$ are two generic probability vectors of size $|T_2|$. $p(X_1 = 0) = x$, $p(X_2 = 1) = y$.

In the rest part, we will discuss concavity condition for genies.

*1) Cardinality bound on Genies:* Traditionally, we need to find some cardinality bounds for the auxiliaries in outer bounds. This is because we have to set the auxiliaries to every possible distribution before we could determine the outer bound from the union of all regions derived. This is not true for the genie case because any particular genie pair corresponds to a valid outer bound. Therefore, we do not necessarily need a cardinality bound. That being said, the essence of having one, though, lies in search of the best genies, *i.e.* to which dimension do we go while we search for the best before we know for sure that there are no better ones beyond. Unfortunately, traditional methods of bounding cardinalities using Caratheodory theorem does not go through as the cardinality bounds for $T_1$ and $T_2$ would end up depending on each other's. We will deploy a tailored method for our case.

By Proposition 4, $g_1(y)$ is concave for $y \in [0, 1]$ if genies satisfy concavity condition. Taking second derivative of $g_1(y)$ with respect to $y$,

$$\frac{d^2 g_1(y)}{dy^2} = \sum_i \left( -\frac{\bar{q}^2(c_i - d_i)^2}{y\bar{q}(c_i - d_i) + d_i} + \frac{pb_i}{y} + \frac{p^2 b_i^2}{\bar{y}pb_i + \bar{p}a_i} \right)$$

$T_2$ is characterized by $\{c_i\}$ and $\{d_i\}$. The following lemma provides cardinality bound for $T_2$.

**Lemma 2.** *Let $n \geq 3$ and $T_{2n}$ be the set of all genies with cardinality $n$. If $T_{2n}(\mathbf{c}, \mathbf{d})$ is a genie defined by*

$\mathbf{c} = (c_1, c_2, \ldots, c_n)$ *and* $\mathbf{d} = (d_1, d_2, \ldots, d_n)$ *such that* $\frac{d^2 g_1(y)}{dy^2} \leq 0$*, then there is always another set of coefficients* $\hat{\mathbf{c}}$*,* $\hat{\mathbf{d}}$ *with* $(n-1)$ *coordinates each such that* $T_{2(n-1)}(\hat{\mathbf{c}}, \hat{\mathbf{d}})$ *defines a genie such that* $\frac{d^2 g_1(y)}{dy^2} \leq 0$*.*

*Proof:*

For $1 \leq i \leq n$, let $\epsilon \geq 0$ and $c_i' = c_i(1+\epsilon l_i)$, $d_i' = d_i(1+\epsilon l_i)$. $\mathbf{c}' = (c_1', \ldots, c_n')$ and $\mathbf{d}' = (d_1', \ldots, d_n')$ form a valid $T_{2n}(\mathbf{c}', \mathbf{d}')$ with some $\mathbf{l} = (l_1, l_2, \ldots, l_n)$ if $\sum_i c_i l_i = 0$, $\sum_i d_i l_i = 0$ and $\epsilon$ small enough. Note that as long as there exists of a non-zero $\mathbf{l}$ independent of $\epsilon$ such that $T_{2n}(\mathbf{c}', \mathbf{d}')$ forcing $\frac{d^2 g_1(y)}{dy^2} \leq 0$ for $0 \leq \epsilon \leq \epsilon_0$, we could increase $\epsilon$ from 0 gradually until for some $i$, $1+\epsilon l_i$ becomes 0. Dropping the 0 coefficients, we get an equivalent genie in $T_{2(n-1)}$. Therefore, it suffices to show the existence of one such $\mathbf{l}$ for $n \geq 3$.

Note that one of the $d_i$'s has to be 0 and the corresponding $c_i$ has to satisfy $\bar{q}c_i \geq p$ in order for $\frac{d^2 g_1(y)}{dy^2}$ to be non-positive when $y \to 0$. In cases where more than one of the $d_i$'s are 0, we could sum over the corresponding $c_i$'s and form a new smart and useful genie with smaller cardinality. Therefore, without loss of generality, we assume that $d_1 = 0$, $\bar{q}c_1 \geq p$ and $d_i > 0, \forall i \geq 2$. All assumptions about $\mathbf{c}$ and $\mathbf{d}$ are as below.

$$\begin{cases} \mathbf{c} \geq \mathbf{0}, \\ \sum_{i=1}^n c_i = 1, \\ \bar{p}c_1 \geq p, \\ d_1 = 0, \\ (d_2, d_3, \ldots, d_n) > \mathbf{0}, \\ \sum_{i=2}^n d_i = 1, \\ -\frac{\bar{q}c_1}{y} + \frac{pb_1}{y} + \frac{p^2 b_1^2}{\bar{y}pb_1 + \bar{p}a_1} \\ +\sum_{i=2}^n \left( -\frac{\bar{q}^2 (c_i - d_i)^2}{y\bar{q}(c_i - d_i) + d_i} + \frac{pb_i}{y} + \frac{p^2 b_i^2}{\bar{y}pb_i + \bar{p}a_i} \right) \leq 0, \forall y \in [0,1]. \end{cases}$$

We need to find $\mathbf{l}$ such that

$$\begin{cases} \mathbf{l} \neq 0, \\ c_1 l_1 + \sum_{i=2}^n l_i c_i = 0, \\ \sum_{i=2}^n l_i d_i = 0, \\ -\frac{\bar{q}c_1(1+\epsilon l_1)}{y} + \frac{pb_1}{y} + \frac{p^2 b_1^2}{\bar{y}pb_1 + \bar{p}a_1} \\ +\sum_{i=2}^n \left( -\frac{\bar{q}^2 (c_i - d_i)^2 (1+\epsilon l_i)}{y\bar{q}(c_i - d_i) + d_i} + \frac{pb_i}{y} + \frac{p^2 b_i^2}{\bar{y}pb_i + \bar{p}a_i} \right) \leq 0, \\ \quad \forall y \in [0,1], \epsilon \in [0, \epsilon_0] \end{cases}$$

Combining above two sets of conditions, and given $\epsilon \geq 0$

$$\begin{cases} \mathbf{l} \neq 0, \\ c_1 l_1 + \sum_{i=2}^n l_i c_i = 0, \\ \sum_{i=2}^n l_i d_i = 0, \\ -\frac{\bar{q}c_1 l_1}{y} - \sum_{i=2}^n \frac{\bar{q}^2 (c_i - d_i)^2 l_i}{y\bar{q}(c_i - d_i) + d_i} \leq 0, \forall y \in [0,1]. \end{cases}$$

Since $c_1 > 0$, set $l_1 = -\frac{\sum_{i=2}^n l_i c_i}{c_1}$. We get the new set of conditions for $l_2, \ldots, l_n$.

$$\begin{cases} \sum_{i=2}^n l_i d_i = 0, \\ \sum_{i=2}^n \frac{l_i d_i (y\bar{q}(c_i - d_i) + c_i)}{y(y\bar{q}(c_i - d_i) + d_i)} \leq 0, \forall y \in [0,1]. \end{cases}$$

Setting $l_i = 0, \forall i \geq 4$, we get

$$\begin{cases} l_2 d_2 + l_3 d_3 = 0, \\ \frac{l_2 d_2 (y\bar{q}(c_2 - d_2) + c_2)}{y\bar{q}(c_2 - d_2) + d_2} + \frac{l_3 d_3 (y\bar{q}(c_3 - d_3) + c_3)}{y\bar{q}(c_3 - d_3) + d_3} \leq 0, \forall y \in [0,1]. \end{cases}$$

Let $l_3 = -\frac{l_2 d_2}{d_3}$. It reduces to show the existence of $(c_2, c_3)$, $(d_2, d_3)$ and $l_2$ such that

$$l_2 d_2 \left( \frac{y\bar{q}(c_2 - d_2) + c_2}{y\bar{q}(c_2 - d_2) + d_2} - \frac{y\bar{q}(c_3 - d_3) + c_3}{y\bar{q}(c_3 - d_3) + d_3} \right) \leq 0, \forall y \in [0,1].$$

This is equivalent to

$$\frac{l_2 d_2 (c_2 d_3 - c_3 d_2)}{(y\bar{q}c_2 + (1 - y\bar{q})d_2)(y\bar{q}c_3 + (1 - y\bar{q})d_3)} \leq 0, \forall y \in [0,1].$$

Therefore, by setting $l_2 = \frac{1}{d_2}$ when $c_2 d_3 \leq c_3 d_2$ and setting $l_2 = -\frac{1}{d_2}$ when $c_2 d_3 > c_3 d_2$, we get a particular non-zero $\mathbf{l}$.

$$\mathbf{l} = \begin{cases} \left( \frac{-c_2 d_3 + d_2 c_3}{c_1 d_2 d_3}, \frac{1}{d_2}, -\frac{1}{d_3}, 0, \ldots, 0 \right), \text{ if } c_2 d_3 \leq c_3 d_2 \\ \left( \frac{c_2 d_3 - d_2 c_3}{c_1 d_2 d_3}, -\frac{1}{d_2}, \frac{1}{d_3}, 0, \ldots, 0 \right), \text{ if } c_2 d_3 > c_3 d_2 \end{cases}$$

■

The above lemma means that for a particular $(p, q)$, the existence of a smart and useful genie with cardinality greater or equal to 3 implies the existence of such a genie within smaller cardinalities. In other words, we could stop searching if we do not find any smart and useful genie within binary choices.

Similar argument can be applied to $T_1$ .

*2) Necessary Conditions:* Based on last section, it is safe to consider only binary genies. Setting $a_1 = a$, $a_2 = \bar{a}$, $b_1 = b$, $b_2 = \bar{b}$, $c_1 = c$, $c_2 = \bar{c}$, $d_1 = d$ and $d_2 = \bar{d}$, we will look at the concavity conditions.

In Proposition 4, we decompose the difference of mutual information as $f(x,y) = (1-x)g_0(y) + xg_1(y)$. Similar for $f(\tilde{x,y})$ defined in the proof of Theorem 4, $\tilde{f}(x,y) = (1-y)\tilde{g}_0(x) + y\tilde{g}_1(x)$. Then by Proposition 4 the concavity condition is equivalent to the condition for $g_1(y)$ to be concave for all $y \in (0,1)$ and $\tilde{g}_1(x)$ is concave for all $x \in (0,1)$.

Take the second derivative of $g_1(y)$ and $\tilde{g}_1(x)$, both has to be non-positive, *i.e.*

$$\sum_{i=1}^2 -\frac{\bar{q}^2 (c_i - d_i)^2}{\bar{y}d_i + y(\bar{q}c_i + pd_i)} + \frac{pb_i}{y} + \frac{p^2 b_i^2}{\bar{y}pb_i + \bar{p}a_i} \leq 0 \qquad (6)$$

and

$$\sum_{i=1}^2 -\frac{\bar{p}^2 (a_i - b_i)^2}{\bar{x}b_i + y(\bar{p}a_i + pb_i)} + \frac{qd_i}{x} + \frac{q^2 d_i^2}{\bar{x}qd_i + \bar{q}c_i} \leq 0 \qquad (7)$$

Note that in (6), either $d_1$ or $d_2$ has to be 0 in order to cancel $\frac{pb_i}{y}$ while $y \to 0^+$. Similarly, either $b_1$ or $b_2$ has to be zero because of (7). Without loss of generosity, we assume that $d_1 = d = 0$ and $b_1 = b = 0$. Therefore, (6) becomes equivalent to, for all $y \in (0,1)$,

$$
\begin{aligned}
&-\frac{\bar{q}c}{y} + \frac{p}{y} - \frac{\bar{q}^2(\bar{c}-1)^2}{\bar{y} + y(\bar{q}\bar{c}+q)} + \frac{p^2}{\bar{y}p + \bar{p}\bar{a}} \leq 0 \\
\Leftrightarrow &\frac{p - \bar{p}c}{y} - \frac{\bar{p}^2c^2}{1 - y\bar{p}c} + \frac{p^2}{\bar{y}p + \bar{p}\bar{a}} \leq 0 \\
\Leftrightarrow &\frac{p}{y} + \frac{p^2}{\bar{y}p + \bar{p}\bar{a}} \leq \frac{\bar{p}c}{y} + \frac{\bar{p}^2c^2}{1 - y\bar{p}c} \\
\Leftrightarrow &\frac{p^2 + p\bar{p}\bar{a}}{\bar{y}p + \bar{p}\bar{a}} \leq \frac{\bar{p}c}{1 - y\bar{p}c} \\
\Leftrightarrow &(p^2 + p\bar{p}\bar{a})(1 - y\bar{p}c) \leq (\bar{p}c)(\bar{y}p + \bar{p}\bar{a}), \forall y \in (0,1) \qquad (8)
\end{aligned}
$$

As the expression is linear in $y$ on both sides, it suffices to check the validity of (8) for when $y = 0$ and $y = 1$, *i.e.* (8) is equivalent to

$$
\begin{cases} p \leq \bar{q}c, \\ p + \frac{p^2}{\bar{p}\bar{a}} \leq \frac{\bar{q}c}{1 - \bar{q}c}. \end{cases}
$$

Rearranging the first inequality we get

$$
\begin{cases} \frac{p}{\bar{p}} \leq \frac{\bar{q}c}{1 - \bar{q}c}, \\ p + \frac{p^2}{\bar{p}\bar{a}} \leq \frac{\bar{q}c}{1 - \bar{q}c}. \end{cases}
$$

Note that $p + \frac{p^2}{\bar{p}\bar{a}} = p(1 + \frac{p/\bar{a}}{\bar{p}}) \geq p(1 + \frac{p}{\bar{p}}) = \frac{p}{\bar{p}}$. Therefore, the first inequality is redundant and we are left with a single constraint

$$
p + \frac{p^2}{\bar{p}\bar{a}} \leq \frac{\bar{q}c}{1 - \bar{q}c}.
$$

Similarly, inequality (7) is equivalent to the following,

$$
q + \frac{q^2}{\bar{q}\bar{c}} \leq \frac{\bar{p}a}{1 - \bar{p}a}.
$$

Further, without loss of generality, we assume $p \leq q$. Putting all the conditions together, we get

$$
\begin{aligned}
0 \leq a \leq 1 \qquad (9) \\
0 \leq c \leq 1 \qquad (10) \\
0 \leq p \leq q \leq 1 \qquad (11) \\
0 \leq p + q \leq 1 \qquad (12) \\
p + \frac{p^2}{\bar{p}\bar{a}} \leq \frac{\bar{q}c}{1 - \bar{q}c} \qquad (13) \\
q + \frac{q^2}{\bar{q}\bar{c}} \leq \frac{\bar{p}a}{1 - \bar{p}a} \qquad (14)
\end{aligned}
$$

Rearranging (13), we have

$$
\bar{p}a \leq \frac{\bar{p}\bar{q}c - p\bar{p}}{\bar{q}c - p^2\bar{q}c - p\bar{p}}
$$

$$
\frac{\bar{p}a}{1 - \bar{p}a} \leq \frac{\bar{q}c - p}{p\bar{q}c}
$$

Note

$$
\frac{\bar{q}c - p}{p\bar{q}c} = \frac{1 - p/\bar{q}c}{p} \leq \frac{\bar{p}}{1 - \bar{p}}
$$

This means (9) is redundant.

Combining with (14) we have the condition

$$
\frac{q\bar{q}\bar{c} + q^2}{\bar{c}} \leq \frac{\bar{q}c - p}{pc}
$$

$$
(1 - pq)\bar{q}c^2 - (1 + p)\bar{q}c + p \leq 0 \qquad (15)
$$

This inequality must holds for some $c \in [0,1]$.

When $c = \frac{1+p}{2(1-pq)}$. $0 \leq c \leq 1$ is given by the following

$$
0 \leq \frac{1+p}{2(1-pq)} = \frac{1+p}{1 + (1 - 2pq)} \leq \frac{1+p}{1 + (1-q)} \leq \frac{1+p}{1 + (1-\bar{p})} = 1
$$

where first inequality is due to $p \leq \frac{1}{2}$ and the second one is due to $q \leq \bar{p}$. So we can let $c = \frac{1+p}{2(1-pq)}$.

Then inequality (15) gives

$$
p - \frac{(1+p)^2\bar{q}}{4(1-pq)} \leq 0
$$

$$
q \leq \frac{1-p}{1+3p}
$$

To satisfy (11), we need $p \leq \frac{1-p}{1+3p}$. That is $0 \leq p \leq \frac{1}{3}$.

Same analysis can be applied to the case $q \leq p$.

Hence we derive the conditions for the existence of smart and useful genie,

$$
\begin{aligned} 0 \leq p \leq \frac{1}{3}, \\ p \leq q \leq \frac{1-p}{1+3p} \end{aligned} \quad \text{or} \quad \begin{aligned} 0 \leq q \leq \frac{1}{3}, \\ q \leq p \leq \frac{1-q}{1+3q} \end{aligned}
$$

## References


[1] V.S. Annapureddy and V.V. Veeravalli. Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region. *Information Theory, IEEE Transactions on*, 55(7):3032 –3050, july 2009.

[2] A. Carleial. Interference channels. *Information Theory, IEEE Transactions on*, 24(1):60–70, 1978.

[3] Hon-Fah Chong, M. Motani, H.K. Garg, and H. El Gamal. On the han-kobayashi region for the interference channel. *Information Theory, IEEE Transactions on*, 54(7):3188 –3195, july 2008.

[4] Abbas El Gamal and Young-Han Kim. *Network Information Theory*. Cambridge University Press, 2012.

[5] Te Han and K. Kobayashi. A new achievable rate region for the interference channel. *Information Theory, IEEE Transactions on*, 27(1):49 – 60, jan 1981.

[6] J Körner and K Marton. General broadcast channels with degraded message sets. *IEEE Trans. Info. Theory*, IT-23:60–64, Jan, 1977.

[7] A.S. Motahari and A.K. Khandani. Capacity bounds for the gaussian interference channel. *Information Theory, IEEE Transactions on*, 55(2):620 –643, feb. 2009.

[8] Xiaohu Shang, G. Kramer, and Biao Chen. A new outer bound and the noisy-interference sum-rate capacity for gaussian interference channels. *Information Theory, IEEE Transactions on*, 55(2):689 –699, feb. 2009.