A new achievable rate region for the 3-user discrete memoryless interference channel

Arun Padakandla, Aria G. Sahebi and S. Sandeep Pradhan
Dept of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109, USA
Email: {arunpr,ariaghs,pradhanv}@umich.edu

***Abstract*—The 3-user discrete memoryless interference channel is considered in this paper. We provide a new inner bound (achievable rate region) to the capacity region for this channel. This inner bound is based on a new class of code ensembles based on asymptotically good nested linear codes. This achievable region is strictly superior to the straightforward extension of Han-Kobayashi rate region from the case of two-users to three-users. This rate region is characterized using single-letter information quantities. We consider examples to illustrate the rate region.**

I. Introduction

In this paper we consider the problem of reliable communication over the discrete memoryless interference channel with 3 users. A schematic of of such a channel is shown in Fig. 1. In this model, 3 transmitters wish to communicate to the corresponding 3 receivers simultaneously over a channel which has three inputs and three outputs. Each output is affected by all the three inputs. This channel was first considered by Carleial in [1] and later in [3]–[5]. The capacity of this channel is still not known even in the case of two-users. It is known in several special cases such as the strong and very strong interference channels. The best known inner bound to the capacity region is given by the Han-Kobayashi achievable rate region [2]. Recently, there has been a lot of activity for the three user channel in special cases such as the additive white Gaussian noise model with and without fading [6]–[10], [13]. For the Gaussian 3-user interference channel it has been shown in [8]–[10] that random lattice codes can provide rates that are beyond what can be obtained using random code ensembles. This approach is called as lattice alignment. We show that even when the channel does not have the additive structure, algebraic structured code ensembles provide superior performance as compared to unstructured code ensembles. In related work, Bresler, Parekh and Tse [11] demonstrate the need for lattice codes in order to achieve within a constant gap of the capacity for the many-to-one interference channels. We build on this approach and consider a framework for the general interference channel. The interested reader is referred to [12], where the authors exploit structure of the interfering signals to provide the largest known rate region for the three user deterministic IC.

We consider a general discrete memoryless case and provide an inner bound to the capacity region for this channel. This is based on our recent work on a new class of algebraic code ensembles called nested linear code ensemble. We build a framework based on reconstructions of bivariate functions. This is similar to the framework developed in [14] for the problem of distributed source coding. In our first step, which is the current paper, we restrict our attention to bivariate functions which are finite fields. In our ongoing work, we are extending this to the case of Abelian groups. In the following, we express all rates in bits, and all logarithms are with base 2.

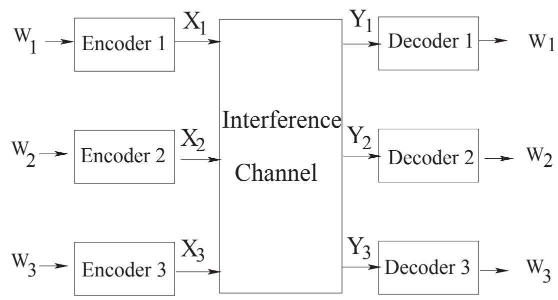

Fig. 1. 3-user interference channel model

II. Preliminaries

A three-user discrete memoryless Interference Channel (DMIC) used without feedback is a septuple $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Y}_3, P_{Y_1,Y_2,Y_3|X_1,X_2,X_3})$ of three input and output alphabets $\mathcal{X}_i$, $\mathcal{Y}_i$ for $i = 1, 2, 3$, and a set of probability distributions $P_{Y_1,Y_2,Y_3|X_1,X_2,X_3}(.|x_1, x_2, x_3)$ on $\mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3$ for all $x_i \in \mathcal{X}_i$, for $i = 1, 2, 3$. The channel is assumed to be memoryless.

Definition II.1. *An (N, M_1, M_2, M_3) transmission system for a given DMIC consists of a mapping for each encoder $e_i : \{1, \ldots, M_i\} \to \mathcal{X}_i^n$ and decoder $g_i : \mathcal{Y}_i^n \to \{1, \ldots, M_i\}$ for $i = 1, 2,$ and 3.*

We assume that the messages (W_1, W_2, W_3) are drawn uniformly from the set $\{1, \ldots, M_1\} \times \{1, \ldots, M_2\} \times \{1, \ldots, M_3\}$. The average error probability of the above transmission system is given by

$$\tau = \frac{1}{M_1 M_2 M_3} \sum_{w_1=1}^{M_1} \sum_{w_2=1}^{M_2} \sum_{w_3=1}^{M_3} \Pr((g_1(Y_1^n), g_2(Y_2^n), g_3(Y_3^n) \neq (w_1, w_2, w_3) | (W_1, W_2, W_3) = (w_1, w_2, w_3).$$

Definition II.2. *A triple (R_1, R_2, R_3) is said to be achievable for a given DMIC if $\forall \epsilon > 0$, there exists an $N(\epsilon)$ such that for all $N > N(\epsilon)$ there exists an (n, M_1, M_2, M_3) transmission systems that satisfies the following conditions*

$$\frac{1}{n} \log M_i \geq R_i - \epsilon, \quad \text{for } i = 1, 2, 3, \quad \tau \leq \epsilon.$$

The set of all achievable rate pairs is the capacity region of the DMIC.

III. Three-to-one IC: First Main Result

A. Coding Theorem

For clarity of exposition, we first consider a special class of interference channels which we call 3-to-1 IC. In this channel, the channel output of the only one of the receivers experiences the interference. We will first describe our new approach to the problem for this class. Then we will consider the more general case. The 3-to-1 IC can be characterized without loss of generality as follows: $P_{Y_1,Y_2,Y_3|X_1,X_2,X_3}$ satisfies the following Markov chains:

$$(X_1, Y_1, X_3, Y_3) \rightarrow X_2 \rightarrow Y_2$$

$$(X_1, Y_1, X_2, Y_2) \rightarrow X_3 \rightarrow Y_3$$

Following is the first main result of this article. For a 3-to-1 interference channel described by $P_{Y_1,Y_2,Y_3|X_1,X_2,X_3}$, let $\mathcal{P}_1$ denote set of probability distributions P_{X_1,U_2,X_2,U_3,X_3} defined over $\mathcal{X}_1 \times \mathcal{U}_2 \times \mathcal{X}_2 \times \mathcal{U}_3 \times \mathcal{X}_3$, having the following property: X_1, (U_2, X_2) and (U_3, X_3) are mutually independent, where $\mathcal{U}_2$, and $\mathcal{U}_3$ are arbitrary finite sets such that $\max_i |\mathcal{X}_i| \leq |\mathcal{U}_2| = |\mathcal{U}_3| = p^r$, where p^r is a prime power. Let $U_1 + U_2$ denote sum of two random variables U_1 and U_2, where addition is with respect to the unique finite field of size p^r. For every such distribution P_{X_1,U_2,X_2,U_3,X_3} let the induced distribution on the set $\mathcal{X}_1 \times \mathcal{U}_2 \times \mathcal{X}_2 \times \mathcal{U}_3 \times \mathcal{X}_3 \times \mathcal{Y}_1 \times \mathcal{Y}_2 \times \mathcal{Y}_3$ be given by

$$P_{X_1,U_2,X_2,U_3,X_3} P_{Y_1,Y_2,Y_3|X_1,X_2,X_3}.$$

Theorem 1. *For every P_{X_1,U_2,X_2,U_3,X_3} in $\mathcal{P}_1$, the following rate region is achievable:*

$$\begin{aligned}
R_1 &\leq \min\{0, H(U_2) - H(U_2 + U_3|Y_1), \\
&\quad H(U_3) - H(U_2 + U_3|Y_1)\} + I(X_1; U_2 + U_3, Y_1) \\
R_2 &\leq I(U_2, X_2; Y_2) \\
R_3 &\leq I(U_3, X_3; Y_3) \\
R_1 + R_2 &\leq I(X_2; Y_2|U_2) + I(X_1; U_2 + U_3, Y_1) + H(U_2) \\
&\quad - H(U_2 + U_3|Y_1) \\
R_1 + R_3 &\leq I(X_3; Y_3|U_3) + I(X_1; U_2 + U_3, Y_1) + H(U_3) \\
&\quad - H(U_2 + U_3|Y_1)
\end{aligned}$$

In the following we give an outline of the coding scheme used to achieve this rate region. We omit the formal proof due to lack of space.

B. Outline of the coding scheme

Note that only the first receiver suffers from the interference. The other two user wish to act as helpers to facilitate partial decoding of the interfering signals at the first receiver. Let U_2 and U_3 denote these parts of the interfering signals X_2, and X_3, respectively, that the first receiver wish to decode. In the standard extension of the Han-Kobayashi scheme from the two to three users, the first receiver decodes the pair (U_2, U_3). However in the approach considered in this paper, we make the first receiver decode a bivariate function of the pair. In particular, we let the first receiver decoder $U_2 + U_3$, where $+$ denotes the addition operation of a finite field. We make the following fundamentally new observation. It turns out that if the codebooks containing information about U_2 and U_3 has the algebraic structure of the corresponding finite field, (i.e., codebook is linear), then *irrespective of the structure of the channel transition probability $P_{Y_1|X_1,X_2,X_3}$*, the first receiver can decode the information $U_2 + U_3$ as long as as the rate of the codebook is not too large. This induces a uniform single-letter distribution on U_2 and U_3. To induce a general non-uniform distribution on U_2 and U_3 we use nested linear codes.

Definition III.1. *Given a finite field $(G, +, \cdot)$, a nested linear code of block-length n is a pair of codes $(\mathcal{C}_1, \mathcal{C}_2)$, where $\mathcal{C}_i$ is a coset of a linear code* [1] *for $i = 1, 2$, with additional constraint being that $\mathcal{C}_2 \subset \mathcal{C}_1$, where a linear code is a vector space over G^n.*

When we use a pair of nested linear codes for a communication problem, we use a carefully chosen set of coset representatives of $\mathcal{C}_2$ in $\mathcal{C}_1$ as the codebook, denoted as $\mathcal{C}_1/\mathcal{C}_2$. Consider the following fact about the asymptotic performance of nested linear codes for point-to-point channel coding that proved in [16].

Fact 1: For any discrete memoryless point-to-point channel $P_{Y|X}$ having input alphabet $\mathcal{X}$ with $|\mathcal{X}| \leq |G|$, there exists a sequence of nested linear codes indexed by the block-length n, such that

$$\lim_{n\rightarrow} \frac{1}{n} \log |\mathcal{C}_1| = \log |G| - H(X|Y)$$

from below,

$$\lim_{n\rightarrow} \frac{1}{n} \log |\mathcal{C}_2| = \log |G| - H(X),$$

from above, and set of coset representatives of $\mathcal{C}_2$ in $\mathcal{C}_1$ forms a good channel code in the usual Shannon sense (probability of decoding error approach 0 as $n \rightarrow \infty$).

We call $\mathcal{C}_1$ as the packing code and $\mathcal{C}_2$ as the shaping code. Now coming back to the 3-to-1 interference channel,

[1] There is a slight abuse of notation here. Strictly speaking, we should call them nested coset codes.

for users 2 and 3, the corresponding information is encoded in pairs (U_2, X_2) and (U_3, X_3), respectively. The two corresponding receivers decode the corresponding pairs. The first receiver first decodes the sum of parts of the interference $U_2 + U_3$, and then decodes the corresponding information bearing signal X_1. We construct a nested linear code for each of these 5 variables. We denote the shaping and packing codes for the variable U_2 as $\mathcal{C}_1(U_1)$ and $\mathcal{C}_2(U_1)$, and similarly for other variables. Let S_i and $S_i + T_i$ denote the rates of the shaping codes, and the packing codes associated with the nested linear codes of U_i for $i = 1, 2$, respectively. Similarly let K_i and $K_i + L_i$ denote the rates of the shaping and packing codes associated with the nested linear code of X_i for $i = 1, 2, 3$. The 5 pairs of nested linear codes are constructed by choosing the elements of their corresponding generator matrices uniformly and independently from the unique finite field of size p^r with replacement, while enforcing the corresponding nested structure. Moreover, we also enforce the condition that $\mathcal{C}_1(U_2) \subset \mathcal{C}_1(U_3)$ or $\mathcal{C}_1(U_3) \subset \mathcal{C}_1(U_2)$, depending on the rate of these codes. It turns out that this is crucial to facilitating the first decoder to recover $U_2 + U_3$. We also choose 5 random dither vectors, where a dither is added to all codewords of a nested linear code. Let R_i denote the transmission rate of the ith encoder. Note that $R_1 = L_1$, $R_2 = L_2 + T_2$ and $R_3 = L_3 + T_3$. Fix the input distribution $P_{X_1,U_2,X_2,U_3,X_3} = P_{X_1} P_{U_2,X_2} P_{U_3,X_3}$.

Encoding: The first encoder, after observing the message W_1, looks for a strongly typical (with respect to P_{X_1}) [15] vector X_1^n from the W_1th coset of $\mathcal{C}_2(X_1)$ in $\mathcal{C}_1(X_1)$, and sends this vector as the channel input. The second encoder partitions its message into two indexes W_{21} and W_{22}, and looks for a pair of vectors (U_2^n, X_2^n) that are strongly typical (with respect to P_{U_2,X_2}) with first vector in the W_{21}th coset of $\mathcal{C}_2(U_2)$ in $\mathcal{C}_1(U_2)$, and first vector in the W_{22}th coset of $\mathcal{C}_2(X_2)$ in $\mathcal{C}_1(X_2)$. Then the encoder sends the vector X_2^n over the channel. Similar encoding is performed at the third encoder. If any encoder is not successful in the corresponding endeavors, it will declare error and sends a random sequence over the channel.

We have shown that the probability of encoding error goes to zero if the rates of shaping codes satisfy the following conditions: for $i = 2, 3$

$$S_i \geq \log p^r - H(U_i), \quad S_i + K_i \geq \log p^{2r} - H(X_i) \tag{1}$$

and

$$K_1 \geq \log p^r - H(X_1). \tag{2}$$

Decoding: The second decoder observes the channel output Y_2^n and looks for a unique sequence pair (U_2^n, X_2^n) with the first vector in $\mathcal{C}_1(U_2)$ and the second vector in $\mathcal{C}_2(X_2)$ such that (U_2^n, X_2^n, Y_2^n) is strongly typical with respect to P_{U_2,X_2,Y_2}. Similar decoding is performed at the third decoder. The first decoder observes Y_1^n and looks for a unique vector $\tilde{U}^n$ in $\mathcal{C}_1(U_1) + \mathcal{C}_1(U_2)$ such that $(\tilde{U}^n, Y_1^n)$ is strongly jointly typical with respect to $P_{U_2+U_3,Y_1}$. If successful, the decoder looks for a unique vector X_1^n in $\mathcal{C}_1(X_1)$ such that $(X_1^n, \tilde{U}^n, Y_1^n)$ is strongly jointly typical with respect to P_{X_1,U_2+U_3,Y_1}. If any decoder is not successful in the corresponding endeavor, it will declare error.

We have shown that the probability of decoding error goes to zero if the rates of packing codes satisfy the following conditions: for $i = 2, 3$

$$K_i + L_i \leq \log p^r - H(X_i|U_i, Y_i) \tag{3}$$

$$S_i + T_i \leq \log p^r - H(U_2 + U_3|Y_1) \tag{4}$$

$$K_i + L_i + S_i + T_i \leq \log p^{2r} - H(U_i, X_i|Y_i) \tag{5}$$

and

$$K_1 + L_1 \leq \log p^{2r} - H(X_1|U_2 + U_3, Y_1) \tag{6}$$

After running a Fourier-Motzkin elimination on equations (1)-(6), we get the rate region presented in the theorem.

IV. General Case: Second Main Result

After going through the basic idea for the case of 3-to-1 interference channel, we are ready to discuss the more general interference channel. We obtain an inner bound to the capacity region that can be characterized using single-letter information quantities. When potentially ever channel output has an interference component, each user may wish to act as a helper to the other two users. We give a brief outline of the coding scheme and omit proof due to lack of space.

Here each user has a pair of auxiliary random variables which will be used to facilitate the other users to decode the interference. The pair (U_{12}, U_{13}) is associated with the first encoder and decoder. Similarly (U_{21}, U_{23}) and (U_{31}, U_{32}) are associated with the second and the third encoder-decoder pairs. The first decoder wishes to recover $U_{21} + U_{31}$ first before decoding its own message contained in (U_{12}, U_{13}, X_1). Similarly $U_{12} + U_{32}$ and $U_{13} + U_{23}$ are recovered first at the second and third decoders respectively. In other words, each encoder has three layers, with the first two being the helper layers. We will use 9 nested linear codes one for each of 6 auxiliary variables and 3 channel input variable. The following is the second main result of this paper.

For a given interference channel described by $P_{Y_1,Y_2,Y_3|X_1,X_2,X_3}$, let $\mathcal{P}_2$ denote the set of all triples of probability distributions P_{U_{12},U_{13},X_1}, P_{U_{21},U_{23},X_2}, and P_{U_{31},U_{32},X_3} defined over $\mathcal{U}_{12} \times \mathcal{U}_{13} \times \mathcal{X}_1$ $\mathcal{U}_{21} \times \mathcal{U}_{23} \times \mathcal{X}_2$ and $\mathcal{U}_{31} \times \mathcal{U}_{32} \times \mathcal{X}_3$. Here $\mathcal{U}_{ij}$ is an arbitrary finite set for $i, j \in \{1, 2, 3\}$ such that $|\mathcal{X}_1| \leq |\mathcal{U}_{21}| = |\mathcal{U}_{31}| = p_1^{r_1}$, $|\mathcal{X}_2| \leq |\mathcal{U}_{12}| = |\mathcal{U}_{32}| = p_2^{r_2}$, and $|\mathcal{X}_3| \leq |\mathcal{U}_{13}| = |\mathcal{U}_{23}| = p_3^{r_3}$, where $p_i^{r_i}$ is a prime power. Let $U_{21} + U_{31}$ denote the sum of the two random variables U_{21} and U_{31}, where addition is with respect to the unique finite field of size $p_1^{r_1}$, and similarly for other cases. For every such triple of distributions, let the induced joint distribution of the auxiliary variables, channel input and channel output be given by

$$P_{U_{12},U_{13},X_1}P_{U_{21},U_{23},X_2}P_{U_{31},U_{32},X_3}P_{Y_1,Y_2,Y_3|X_1,X_2,X_3}.$$

We will use the following notation for a compact description of the rate region. We will use double indexed rates such as S_{ij} and double indexed random variables such as U_{ij}, where each index can take values in $I \triangleq \{1,2,3\}$ and $i \neq j$. For a given pair (i,j) let k denote the element in I such that $k \neq i$ and $k \neq j$. For example when $(i,j) = (1,3)$, $k = 2$. Let $U_{i\bar{i}}$ denote the collection of random variables $\{U_{ij}, U_{ik}\}$. Let $S_{i\bar{i}}$ denote the sum rate $S_{ij} + S_{ik}$. For example $S_{1\bar{1}} = S_{13} + S_{12}$. Let $\tilde{U}_i$ denote the sum $U_{ji} + U_{ki}$, where $+$ is the addition operation of the corresponding finite field. Let $\tilde{T}_i$ denote $\max\{T_{ji}, T_{ki}\}$. For example $\tilde{T}_1 = \max\{T_{21}, T_{31}\}$. Observe that in $\tilde{}$ notation, the index i becomes the second index. Let $A_i = \log p_i^{r_i}$, and let $A_{\bar{i}} = A_j + A_k$, and $A_I = A_1 + A_2 + A_3$.

Theorem 2. *For every triple of distributions in $\mathcal{P}_2$, the following rate region is achievable: for all (i,j) such that $i, j \in \{1,2,3\}$ and $i \neq j$: $R_i = L_i - K_i + T_{i\bar{i}} - S_{i\bar{i}}$, $T_{ij} \geq S_{ij}$, $L_i \geq K_i$, $S_{ij} \leq 1$ and $K_i \leq 1$,*

$$\begin{aligned}
S_{ij} &\geq A_j - H(U_{ij}) \\
S_{i\bar{i}} &\geq A_{\bar{i}} - H(U_{i\bar{i}}) \\
S_{i\bar{i}} + K_i &\geq A_I - H(X_i) \\
L_i &\leq A_i - H(X_i|U_{i\bar{i}}, \tilde{U}_i, Y_i) \\
L_i + T_{ij} &\leq A_{\bar{k}} - H(X_i, U_{ij}|U_{ik}, \tilde{U}_i, Y_i) \\
L_i + T_{i\bar{i}} &\leq A_I - H(X_i, U_{i\bar{i}}|\tilde{U}_i, Y_i) \\
L_i + \tilde{T}_i &\leq 2A_i - H(X_i, \tilde{U}_i|U_{i\bar{i}}, Y_i) \\
L_i + T_{ij} + \tilde{T}_i &\leq 2A_i + A_j - H(X_i, U_{ij}, \tilde{U}_i|U_{ik}, Y_i) \\
L_i + T_{i\bar{i}} + \tilde{T}_i &\leq A_I + A_i - H(X_1, U_{i\bar{i}}, \tilde{U}_i|Y_i).
\end{aligned}$$

Remark: It is not yet clear to us whether this is amenable to Fourier-Motzkin elimination. This rate region has a structure that is similar to polymatroids and could potentially be exploited in Fourier-Motzkin elimination for a compact description. The first 3 equations are shaping constraints, and the next 6 are packing constraints. The above compact description has 12 bounds related to shaping codes, 36 bounds related to packing codes, and 15 non-negativity bounds.

V. Example

In this section, we provide a simple example for which the rate region achievable using nested linear codes is strictly larger than that achievable using the natural extension of Han-Kobayashi region for three users.

Consider the binary 3-to-1 interference channel depicted in Figure 2 whose input output relationship is described through the equations $Y_1 = X_1 \oplus X_2 \oplus X_3 \oplus N_1$, $Y_2 = X_2 \oplus N_2$ (depicted as W_2 in the figure) and $Y_3 = X_3 \oplus N_3$ (depicted as W_3 in the figure), where $\oplus$ denotes modulo-2 sum, N_1, N_2 and N_3 are mutually independent Bernoulli random variables with $P(N_1 = 1) = \delta$, $P(N_2 = 1) = P(N_3 = 1) = \epsilon$. We assume an average Hamming weight constraint q on the input of user 1. i.e., $\frac{1}{n}\mathbb{E}\left\{\sum_{i=1}^{n} 1_{\{X_{1i}=1\}}\right\} \leq q$.

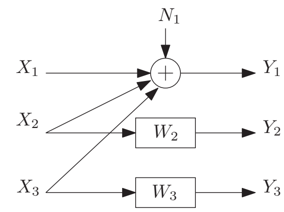

Fig. 2. A simple many to one interference channel.

We now state the rate region achievable for this 3-to-1 IC using structured codes. Consider the rate region described in theorem 1. Let X_1, X_2, X_3 be mutually independent with $P(X_1 = 1) = q$, $P(X_2 = 1) = P(X_3 = 1) = \frac{1}{2}$. Let $U_2 = X_2$, $U_3 = X_3$. It can be verified that if $q * \delta < \epsilon$, then $(h_b(q*\delta) - h_b(\delta), 1 - h_b(\epsilon), 1 - h_b(\epsilon))$ is achievable, where $h_b(q) = -q\log_2 q - (1-q)\log_2(1-q)$ is the binary entropy function and $q*\delta = \delta(1-q) + q(1-\delta)$. Indeed, we employ the same linear code of rate $1 - h_b(\epsilon)$ for users 2 and 3. Since $1 - h_b(\epsilon)) < 1 - h_b(q*\delta)$, user 1 can decode $X_2^n + X_3^n$, the sum of the codewords transmitted by users 2 and 3. After subtracting off the interference, decoder 1 decodes user 1's message at rate $h_b(q*\delta) - h_b(\delta)$. Therefore, each transmitter receiver pair communicates at their respective point to point capacities as if there were no interference.

We now prove the rate region achievable using the natural extension of Han-Kobayashi is strictly sub optimal. We begin with a characterization of Han-Kobayashi region for 3-to-1 IC. A rate triple (R_1, R_2, R_3) is achievable using the Han-Kobayashi scheme if there exists a distribution $p_{X_1U_2X_2U_3X_3Y_1Y_2Y_3} = p_{X_1}p_{U_2X_2}p_{U_3X_3}p_{Y_1Y_2Y_3|X_1X_2X_3}$ such that

$$\begin{aligned}
&R_1 \leq \min\{I(X_1;Y_1|U_2U_3), I(X_1U_3;Y_1|U_2), I(X_1U_2;Y_1|U_3)\} \\
&R_j \leq I(X_j;Y_j) : j = 2,3 \\
&R_1 + R_2 \leq I(X_2;Y_2|U_2) + I(X_1U_2;Y_1|U_3) \\
&R_1 + R_3 \leq I(X_3;Y_3|U_3) + I(X_1U_3;Y_1|U_2) \\
&R_1 + R_2 + R_3 \leq I(X_1U_2U_3;Y_1) + \sum_{j=2}^{3} I(X_j;Y_j|U_j). \qquad (7)
\end{aligned}$$

Lemmas V.1 enables us in proving strict suboptimality of the above rate region.

Lemma V.1. *For any collection $(U_2, X_2, U_3, X_3, X_1, Y_1)$ of random variables that satisfy, (i) X_1, X_2, X_3 are binary and mutually independent, (ii) $p_{U_2X_2U_3X_3X_1Y_1} = p_{U_2X_2}p_{U_3X_3}p_{X_1}1_{\{Y_1 = X_1 \oplus X_2 \oplus X_3 \oplus N_1\}}$ where X_1, X_2, X_3, N_1 ,are mutually independent and $P(N_1 = 1) = \delta$, and (iii) $H(X_2 \oplus X_3|U_2, U_3) > 0$, we have $I(X_1;Y_1|U_2,U_3) < I(X_1;Y_1|X_2 \oplus X_3)$.*

Proof: The proof is based on a simple observation that if $Z_j : j = 1, 2, 3$ are binary random variables such that (Z_1, Z_2) is independent of Z_3, then $|H(Z_1 \oplus Z_3) - H(Z_2 \oplus Z_3)| \leq |H(Z_1) - H(Z_2)|$ with the inequality being strict if $H(Z_3) > 0$. This can be proved by using the strict concavity of the binary entropy function and is omitted.

We now prove the lemma.

$$
\begin{aligned}
&I(X_1; Y_1|U_2U_3) = H(Y_1|U_2U_3) - H(Y_1|X_1U_2U_3) \\
&= \sum_{u_2,u_3} p_{U_2U_3}(u_2, u_3) H(Y_1|U_2 = u_2, U_3 = u_3) \\
&\quad - \sum_{x_1,u_2,u_3} p_{X_1U_2U_3}(x_1, u_2, u_3) H(Y_1|^{X_1=x_1, U_2=u_2,}_{U_3=u_3}) \\
&= \sum_{u_2,u_3} p_{U_2U_3}(u_2, u_3) H\left(\substack{(X_1 \oplus N_1) \\ \oplus \\ (X_2 \oplus X_3)} | U_2 = u_2, U_3 = u_3\right) \\
&\quad - \sum_{x_1,u_2,u_3} p_{X_1U_2U_3}(x_1, u_2, u_3) H\left(\substack{(x_1 \oplus N_1) \\ \oplus \\ (X_2 \oplus X_3)} |^{X_1=x_1, U_2=u_2,}_{U_3=u_3}\right) \\
&= \sum_{u_2,u_3} p_{U_2U_3}(u_2, u_3) H\left(\substack{(X_1 \oplus N_1) \\ \oplus \\ (X_2 \oplus X_3)} | U_2 = u_2, U_3 = u_3\right) \\
&\quad - \sum_{x_1,u_2,u_3} p_{X_1U_2U_3}(x_1, u_2, u_3) H\left(\substack{(N_1) \\ \oplus \\ (X_2 \oplus X_3)} |^{U_2=u_2,}_{U_3=u_3}\right) \\
&= \sum_{u_2,u_3} p_{U_2U_3}(u_2, u_3) H\left(\substack{(X_1 \oplus N_1) \\ \oplus \\ (X_2 \oplus X_3)} | U_2 = u_2, U_3 = u_3\right) \\
&\quad - \sum_{u_2,u_3} p_{U_2U_3}(u_2, u_3) H\left(\substack{N_1 \\ \oplus \\ (X_2 \oplus X_3)} |^{U_2=u_2,}_{U_3=u_3}\right) \\
&\leq \sum_{u_2,u_3} p_{U_2U_3}(u_2, u_3) H\left((X_1 \oplus N_1)|U_2 = u_2, U_3 = u_3\right) \\
&\quad - \sum_{u_2,u_3} p_{U_2U_3}(u_2, u_3) H\left(N_1|U_2 = u_2, U_3 = u_3\right) \qquad (8) \\
&= H\left((X_1 \oplus N_1)|U_2, U_3\right) - H\left(N_1|U_2, U_3\right) \\
&= H\left(X_1 \oplus N_1\right) - H\left(N_1\right) = I(X_1; Y_1|X_2 \oplus X_3),
\end{aligned}
$$

where (8) follows from substituting $p_{X_1 \oplus N_1|U_2U_3}(\cdot|u_2, u_3)$ for p_{Z_1}, $p_{N_1|U_2U_3}(\cdot|u_2, u_3)$ for p_{Z_2} and $p_{X_2 \oplus X_3|U_2U_3}(\cdot|u_2, u_3)$ for p_{Z_3} in the above observation. ■

The import of lemma V.1 is that $I(X_1; Y_1|U_2, U_3) = I(X_1; Y_1|X_2 \oplus X_3)$ only if $H(X_2 \oplus X_3|U_2, U_3) = 0$. This implies $H(X_j|U_j) = 0$ for $j = 1, 2$. Substituting this for the sum bound in (7), we obtain

$$
\begin{aligned}
R_1 + R_2 + R_3 &\leq I(X_1U_2U_3; Y_1) + \sum_{j=2}^{3} I(X_j; Y_j|U_j) \\
&\leq 1 - h_b(\delta) + 0 + 0 = 1 - h_b(\delta)
\end{aligned}
$$

The sum rate achievable using linear codes is $h_b(q * \delta) - h_b(\delta) + 2 - 2h_b(\epsilon) < h_b(q * \delta) - h_b(\delta) + 2 - 2h_b(q * \delta) = (1 - h_b(\delta)) + (1 - h_b(q * \delta))$, where the first inequality follows from our assumption $q * \delta < \epsilon$. Clearly, the sum rate achievable using linear codes is strictly larger than that achievable using the Han-Kobayashi scheme.

This example illustrates our coding technique in it's simplest form. Our main contribution lies in being able to generalize this idea to arbitrary 3-to-1 and general interference channels. For example, if W_2 and W_3 are asymmetric channels, then linear codes do not achieve their respective capacities. The ensemble of nested linear codes achieve capacity of asymmetric channels without compromising linearity and can therefore be employed to obtain interference alignment at decoder 1.

Acknowledgment

The authors would like to thank Viveck Cadambe for several discussions on the topic of interference channels. This work was supported by NSF grants CCF-0915619 and CCF-1116021.

References

[1] A. B. Carleial, "A case where interference does not reduce capacity," IEEE Trans. on Inform. Theory, vol. 21, no. 5, pp. 569-570, Sept. 1975.
[2] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," IEEE Trans. Inform. Theory, vol. 27, no. 1, pp. 49-60, Jan. 1981.
[3] H. Sato, "The capacity of the Gaussian interference channel under strong interference," IEEE Trans. Inform. Theory, vol. 27, no. 6, pp. 786-788, Nov. 1981.
[4] M. Costa, "On the Gaussian Interference Channel," IEEE Trans. Inform. Theory, vol. 31, no. 5, pp. 607–615, Sep. 1985.
[5] M. H. M. Costa and A. A. E. Gamal "The Capacity Region of the Discrete Memoryless Interference Channel with Strong Interference" IEEE Trans. Inform. Theory, vol. 33, no. 5, pp. 710-711, Sept. 1987.
[6] I. Sason "On Achievable Rate Regions for the Gaussian Interference Channel" IEEE Trans. Inform. Theory, vol. 50, no. 6, pp. 1345-1356, June 2004.
[7] V. S. Annapureddy and V. Veeravalli, "Gaussian interference networks: sum capacity in the low interference regime and new outer bounds on the capacity region," Feb. 2008, submitted to IEEE Trans. Inform. Theory.
[8] A. Jafarian and S. Vishwanath, "Achievable Rates for Gaussian Interference Channels", *IEEE Trans. on Inf. Th.*, October 2010.
[9] S. Shridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar, "Lattice Coding for K User Gaussian Interference Channels", *IEEE Trans. on Inf. Th.*, April 2010.
[10] A. Jafarian, S. Vishwanath, "Gaussian Interference Networks: Lattice Alignment", *IEEE Inf. Th. Workshop*, January 2010.
[11] Bresler, G.; Parekh, A.; Tse, D.N.C.; , "The Approximate Capacity of the Many-to-One and One-to-Many Gaussian Interference Channels," Inf. Th., IEEE Transactions on , vol.56, no.9, pp.4566-4592, Sept. 2010 doi: 10.1109/TIT.2010.2054590.
[12] Bandemer, B.; El Gamal, A.; , "An achievable rate region for the 3-user-pair deterministic interference channel," Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on , vol., no., pp.38-44, 28-30 Sept. 2011 doi: 10.1109/Allerton.2011.6120147.
[13] V. Cadambe, Syed A. Jafar, C. Wang, "Interference Alignment with Asymmetric Complex Signaling – Settling the Host-Madsen-Nosratinia Conjecture", *IEEE Trans. on Inf. Th.*, Sep. 2010, Vol. 56, Issue: 9, Pages: 4552-4565.
[14] D. Krithivasan and S. S. Pradhan, "distributed source coding using abelian group codes: A new achievable rate-distortion function", IEEE Trans. on Inform. Theory, vol. 57, March 2011.
[15] I. Csiszár and J. Korner, *Inf. Th.: Coding Theorems for Discrete Memoryless Systems.* Academic Press Inc. Ltd., 1981.
[16] A. Padakandla and S. S. Pradhan, "Nested linear codes achieve Marton's inner bound for general broadcast channels", ISIT 2011, St. Petersburg, Russia.